

1 GCGGGGCGGA ACAGATCGCA GACCTGGGGG TTCGCAGAGC CGCCAGTGGG
51 GAGATGTTGA AGTTCAAATA TGGAGCGCGG AATCCTTTGG ATGCTGTGTC
101 TGCTGAACCC ATTGCCAGCC GGGCCTCCAG GCTGAATCTG TTCCTCCAGG
151 GGAAACCACC CTTTATGACT CAACAGCAGA TGCTCTCTCT TTCCTCGAGAA
201 GGGATATTAG ATGCCCTCTT TGTTCTCTTT GAAGAATGCA GTCAGCCTGC
251 TCTGATGAAG ATTAAGCACG TGAGCAACTT TGTCCGGAAG TATTCGGACA
301 CCATAGCTGA GTTACAGGAG CTCAGCCTT CGGCAAGGA CTTCGAAGTC
351 AGAAGTCTG TAGGTTGTGG TCACCTTGTG GAAGTGCAGG TGGTAAGAGA
401 GAAAGCAACC GGGGACATCT ATGCTATGAA AGTGATGAAG AAGAAAGGCTT
451 TATTGGCCCA GGAGCAGGTT TCATTTTTTG AGGAAGAGCG GAACATATTA
501 TCTCGAAGCA CAAGCCCGTG GATCCCCCAA TTACAGTATG CCTTTCAGGA
551 CAAAAATCAC CTTTATCTGG TCATGGAATA TCAGCCTGGA GGGGACTTGC
601 TGTCACCTTT GAATAGATAT GAGGACCAGT TAGATGAAAA CCTGATACAG
651 TTTTACCTAG CTGAGCTGAT TTTGGCTGTT CACAGCGTTC ATCTGATGGG
701 ATACGTGCAT CGAGACATCA AGCCTGAGAA CATTCTCGTT GACCGCACAG
751 GACACATCAA GCTGGTGGAT TTTGGATCTG CCGCGAAAAAT GAATTCAAAC
801 AAGATGGTAA AAAATGGAAT AAGATAGCTT AATAGAGTTT ATACTAAAAA
851 GTGTTCTTGG TCCTCCTAAG TTTGGGAAGT GTTGGGATAA AATGGTGAAC
901 AATGTTTGG AGCCTTTGGC AGTGTATGGG GGTGGGGACA GGGACACAGA
951 ACCATTTCCT AGACCGTGGC ACCTTTTAT TTATAGTGCC TGTTAATACC
1001 CTCCAAGACA TTTTATGGAG CATTGTTATA GTTTGTTAG AAATAAAGGA
1051 AATGCTTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1101 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-53
Start Codon: 54
Stop Codon: 825
3'UTR: 828

FIG. 1A

Homologous proteins:

Top 10 BLAST Hits

	Score	E
CRA 18000005161385 /altid=gi 3360512 /def=gb AAC27932.1 (AF070...	465	e-130
CRA 18000005168811 /altid=gi 3599509 /def=gb AAC72823.1 (AF086...	464	e-129
CRA 18000005168810 /altid=gi 3599507 /def=gb AAC72822.1 (AF086...	464	e-129
CRA 89000000196974 /altid=gi 7294566 /def=gb AAF49906.1 (AE003...	213	4e-54
CRA 84000015363786 /altid=gi 13648270 /def=ref XP_008814.3 Rho...	209	5e-53
CRA 67000040980049 /altid=gi 13592049 /def=ref NP_112360.1 Rho...	209	5e-53
CRA 18000005044861 /altid=gi 6677759 /def=ref NP_033097.1 Rho...	209	5e-53
CRA 18000005028208 /altid=gi 4885583 /def=ref NP_005397.1 Rho...	209	5e-53
CRA 18000005169610 /altid=gi 3628755 /def=gb AAC36189.1 (U4242...	209	5e-53
CRA 18000005236627 /altid=gi 5174413 /def=ref NP_006026.1 CDC4...	206	3e-52

BLAST hits to dbEST:

	Score	E
gi 14343911 /dataset=dbest /taxon=960...	831	0.0
gi 11292270 /dataset=dbest /taxon=96...	743	0.0
gi 12111020 /dataset=dbest /taxon=96...	517	e-144
gi 12362084 /dataset=dbest /taxon=96...	492	e-136

EXPRESSION INFORMATION FOR MODULATORY USE:

library source of BLAST dbEST hits:

gi|14343911 blood, Proliferating Human Erythroid Cells
gi|11292270 brain, glioblastoma with EGFR amplification

Tissue source of cDNA clone:

whole liver

FIG. 1B

1 MLKFKYGARN PLDAGAAEPI ASRASRLNLF FQKPPFMTQ QQMSPLSREG
51 ILDALFVLFE ECSQPALMKI KHSVNFVRKY SDTIAELQEL QPSAKDFEVR
101 SLVGCGHFAE VQVVREKATG DIYAMKVMKK KALLAQEQVS FFEEERNILS
151 RSTSPWIPQL QYAFQDKNHL YLVMEYQPGG DLLSLNRYE DQLDENLIQF
201 YLAELILAVH SVHLMGYVHR DIKPENILVD RTGHIKLVDF GSAAKMNSNK
251 MVKNGIR (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

Prosite results:

PDOC00004 PS00004 CAMP_PHOSPHO_SITE

CAMP- and cGMP-dependent protein kinase phosphorylation site

78-81 RKYS

PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 2

1 93-95 SAK

2 248-250 SNK

PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 3

1 83-86 TIAE

2 93-96 SAKD

3 140-143 SFFE

FIG. 2A

PD0C00008 PS00008 MYRISTYL
N-myristoylation site
50-55 GILDAL

PD0C00100 PS00107 PROTEIN_KINASE_ATP
Protein kinases ATP-binding region signature
103-126 VCGHFQAEVQVREKATGDIYAMK

PD0C00100 PS00108 PROTEIN_KINASE_ST
Serine/Threonine protein kinases active-site signature
217-229 YVHRDIKPENILV

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	197	217	0.789	Putative

FIG. 2B

BLAST Alignment to Top Hit:

>CRA|18000005161385 /altid=gi|3360512 /def=gb|AAC27932.1| (AF070065)
Citron-k kinase [Rattus norvegicus] /org=Rattus
norvegicus /taxon=10116 /dataset=nraa /length=448
Length = 448

Score = 465 bits (1183), Expect = e-130
Identities = 228/251 (90%), Positives = 240/251 (94%)
Frame = +3

Query: 54 MLKFKYGARNPLDAGAAEPIASRASRLNLFQCKPPFMTQQQMSPLSREGILDALFVLFE 233
MLKFKYG RNP +A A EPIASRASRLNLFQCKPP MTQQQMS LSREG+LDALFVL E
Sbjct: 1 MLKFKYGVVRNPSEASAPEPIASRASRLNLFQCKPPLMTQQQMSALSREGVLDALFVLL E 60

Query: 234 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 413
ECSQPALMKIKHVS+FVRKYSDTIAEL+ELQPS +DFEVRSLVCGGHFAEVQVVRKATG
Sbjct: 61 ECSQPALMKIKHVSFVRKYSDTIAELRELQPSVRDFEVRSLVCGGHFAEVQVVRKATG 120

Query: 414 DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLM EYQPGG 593
D+YAMK+MKK AL AQEQVSFFEEERNILS+STSPWIPQLQYAFQDKN+LYLVM EYQPGG
Sbjct: 121 DVYAMKIMKKKALRAQEQVSFFEEERNILSQSTSPWIPQLQYAFQDKNNLYLVM EYQPGG 180

Query: 594 DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD F 773
DLLSLLNRYEDQLDEN+IQFYLAELILAVHSVH MGYVHRDIKPENIL+DRTGHIKLVD F
Sbjct: 181 DLLSLLNRYEDQLDENMIQFYLAELILAVHSVHMQMGYVHRDIKPENILIDRTGHIKLVD F 240

Query: 774 GSAAKMNSNKM 806
GSAAKMNSNK+

Sbjct: 241 GSAAKMNSNKV 251 (SEQ ID NO:4)

FIG. 2C

>CRA|18000005168811 /altid=gi|3599509 /def=gb|AAC72823.1| (AF086824)
rho/rac-interacting citron kinase [Mus musculus]
/org=Mus musculus /taxon=10090 /dataset=nraa
/length=2055
Length = 2055

Score = 464 bits (1180), Expect = e-129
Identities = 227/251 (90%), Positives = 242/251 (95%)
Frame = +3

Query: 54	MLKFKYGARNPLDAGAAEPIASRASRLNLFQKGKPPFMTQQQMSPLSREGILDALFVLFE	233
	MLKFKYG RNP +A A+EPIASRASRLNLFQKGKPP MTQQQMS LSREG+LDALF LFE	
Sbjct: 1	MLKFKYGVVRNPPEASASEPIASRASRLNLFQKGKPPPLMTQQQMSALSREGMLDALFALFE	60
Query: 234	ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATG	413
	ECSQPALMK+KHVS+FV+KYSDTIAEL+ELQPSA+DFEVRSLVGCGHFAEVQVVREKATG	
Sbjct: 61	ECSQPALMKMKHVSSFVQKYSDTIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATG	120
Query: 414	DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMVEYQPGG	593
	D+YAMK+MKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKN+LYLVMVEYQPGG	
Sbjct: 121	DVYAMKIMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNNLYLVMVEYQPGG	180
Query: 594	DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRGTGHIKLVDF	773
	D LSLNRYEDQLDE++IQFYLAELILAVHSVH MGYVHRDIKPENIL+DRGTG IKLVDF	
Sbjct: 181	DFLSLLNRYEDQLDESMIQFYLAELILAVHSVHQMGYVHRDIKPENILIDRTGTEIKLVDF	240
Query: 774	GSAAKMNSNKM 806	
	GSAAKMNSNK+	
Sbjct: 241	GSAAKMNSNKV 251 (SEQ ID NO:5)	

FIG. 2D

```
>CRA|18000005168810 /altid=gi|3599507 /def=gb|AAC72822.1| (AF086823)
rho/rac-interacting citron kinase short isoform [Mus
musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
/length=494
Length = 494

Score = 464 bits (1180), Expect = e-129
Identities = 227/251 (90%), Positives = 242/251 (95%)
Frame = +3

Query: 54 MLKFKYGARNPLDAGAAEPIASRASRLNLFQKGKPPFMTQQQMSPLSREGILDALFVLFE 233
MLKFKYG RNP +A A+EPIASRASRLNLFQKGKPP MTQQQMS LSREG+LDALF LFE
Sbjct: 1 MLKFKYGVVRNPPEASASEPIASRASRLNLFQKGKPPPLMTQQQMSALSREGMLDALFALFE 60

Query: 234 ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATG 413
ECSQPALMK+KHVS+FV+KYSDTIAEL+ELQPSA+DFEVRSLVGCGHFAEVQVVREKATG
Sbjct: 61 ECSQPALMKMKHVSSFVQKYSDTIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATG 120

Query: 414 DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPGG 593
D+YAMK+MKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKN+LYLVMEYQPGG
Sbjct: 121 DVYAMKIMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNNLYLVMEYQPGG 180

Query: 594 DLLSLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDF 773
D LSLNRYEDQLDE++IQFYLAELILAVHSVH MGYVHRDIKPENIL+DRTG IKLVDF
Sbjct: 181 DFLSLNRYEDQLDESMIQFYLAELILAVHSVHQMGYVHRDIKPENILIDRTGEIKLVDF 240

Query: 774 GSAAKMNSNKM 806
GSAAKMNSNK+
Sbjct: 241 GSAAKMNSNKV 251 (SEQ ID NO:6)
```

FIG. 2E

Hmmer search results (Pfam):
Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00069	Eukaryotic protein kinase domain	131.2	1.9e-35	1
CE00359	E00359 bone_morphogenetic_protein_receptor	5.8	0.57	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00359	1/1	219	249 ..	274	304 ..	5.8	0.57
PF00069	1/1	97	256 ..	1	147 [.	131.2	1.9e-35

FIG. 2F


```
1 GGGTGACGGA GTGAGATTCT GTCTAAGAAA AAAGAAAAAA AAAGAGGTGC
51 TTGATAAATA GTAGCTATCC ATTATTGGCC CCGGGAACAA GAAGTAAGTT
101 ATGTTTGGGG AAGGAAAAAA GAACAAATGT GTATTAAGCA AGCCTGTAGC
151 TCTAATTATG TGCTGGTGTG CGTGTGTGTG TGTGTGTGTG TGAGAGAGAG
201 AACACATCTC CAGTTCTGTC TACTGTAGAA TTAGGAGAGT AAAAAAGGA
251 CTTTACATAT ATAAATAGAA CACACACACA CACACATGCG TGCACACATA
301 TACACACAAT TTAATCATTA TGAAACCACA TCCATATTGT TGCTACCTAG
351 GTTAAGAAAT AGATCACAGC AGCACCCCAA CACCCTGAAA GGCCTCCATC
401 CCAACCCAG GTAACCTACTA TTCTGGCTGT TGCTTTCTTT ATGGTTTTGT
451 CATTACTTTA AACAATGACA AAAACTGCAA TGATTTCAT CAACCTAATA
501 CATCCCTCCT TAAACAATGT TGCTTTGTTT TGTCCTGTTT TGGAACCTAT
551 AAGAAATGGAA TCATAATGGA ATCATATGTT ATTTTCTTGC TTCCTTCATT
601 AGGCCTTGTT TTGAGACTCA TTATGTCATT GTGGTTAGTT GCAGTTTATT
651 CTTTTTCATT GCTTGTGAAA ACACTGCAAT ATACAATTTT GTCTTTTCTA
701 CTGCTGATGG ACATTTATAT CACTTCCAGT TTTTTCGAA CACTATTTTG
751 TATTCTTATA CACATCTCTT GGTGTACATA AGTAGGAGTT TCTCGCCGGC
801 GTGGTGGCTC AGGGCCTGTA ATCTCAGCAC TTTGGGAGGC CGAGGTGGGC
851 AGATCACTCG AGGTCAGGAG TTCAAGACCA GCCTGGCCAA CACGGTGAAA
901 CCCCATCTCT ACTAAAAATA CAAACAATTG GGCATGGTGG CATGCACCTG
951 TAATCCAGT TACTTGGGAG GATGAGACAA GAGAATAGCT TGAACCTGGG
1001 AGGTGGAGGT TGCAGTGAGG CGAGATCGTG CCATTGCACT CCAGCCTGGG
1051 AGACAGAGCA AGACTCCATC TCAAATAAA TAAATAAATA GGAGTTTTTC
1101 TTAGGTAGAG AACTACACC TAGCAATAGT CATAGAATGC ACAAATCTTC
1151 AATGTTAGCA AATAATGCCA AACTTTTTTT TCAAATTTCA AAGAGATTGT
1201 ATCCATTTAC ACGCCTACGG GTACTGTATA AGTGTGTGTA CTCCACATC
1251 TTCGAAACA CTGTCACATC CTTTTGTTGT TGTGTTCTC GAATTTGAGT
1301 GTTATTCTTT CTCACTGTGA CTTTATTTTT CATATTTTCT GATTATGAAC
1351 GAGGTTGACA ACTTTCACAC ATTTGTTGGT CATCTGGATT TCCTTTTTGG
1401 TGAAGTGCCT GTTTAAGTAT CTCGTCTATA ATTTATTTTA AAGTGTCCCT
1451 TCAGACAGTC TCAATGACTG TCACCAACTC CTTGCAGGGC AGTCAGCCCG
1501 GAGATAGAGT AATCAAGGTA GGTGAAGTC AAGCTCAAAA CATTGCTGTC
1551 CTCAGCTGTA GCAGAGGACC ACTGGGCTTC CCCAGGTAAC AAGTACTTCT
1601 ACCTTAGCCA CATGAGAGAG AAAGAAGACC AGGCAGAGCA GCCTGGCTGC
1651 CTTCTCTCTT GCAGGTGGCC GAGAGCAGGG GACAGCGCCC TGGCGACCTC
1701 CTCAGGGATC CTAGATTAAC AGTCGCGTCC TCAAACGCAG CATCCTGCGT
1751 AACCGCCAAT TTCAAACCTC CAAGACCTGC CCTGCTGATT TTGCCCTTCC
1801 CTTTTTCCCG TTGGTCGCGA GTCAAAGGAA GATGCAATTT GATTGGCTCT
1851 CCCCTTCACT TTCCTCCATG CCTTTAGGGA CATGGGCGGG GCCTGGCTGA
1901 GACGCCCATG TCTATCATAG GAGCGGAGAC GCTGATTGGT CCAAACACGG
1951 CTGAGACCCG CCCGCGCCGT TCCTCGGGTT CAAACGCGGC GCGGGGAGGC
2001 GCGGGGCGGA ACAGATCGCA GACCTGGGGG TTCGAGAGC GTGAGTCTGA
2051 TCCCCAGAC CCAATTCTAC CGCACCCGGC TCTGCAAGGC CAGGGGAGGG
2101 CCGCTCCAC CCATACAAGT CCCGGGTTTC CCTCCGCCC CGGGGAGGGC
2151 GGCGATTCCA CCCCAGGGC TGCGGGAGGC CTGGAGGGTC TTCCGGGGCT
2201 AGCTGTGCGC GCGCCACCT TCCTTGGGAG CCGAGGGGTC AGCCGAGTGG
2251 TGCTGGGGCA GGAGGCTTGC TCCTCCCCTA AACCAGGCGG AGTGCTTTGT
```

FIG. 3A

```
2301 CTCTTCAGCT CTGCCTCCTG TCAGCACTAA CTGCATTATT CTGCCCAGTG
2351 TAGTCGGCCG GTTCCTTATT ATCTGCGTGA ACTTAGCCAT TTACTTAACC
2401 TCTCTGTTTC AGCGTATTCA TACCCCGTGC CCACCCCATC ACCTCATGAT
2451 GCCCCCGCCT CTTTCGCTCT GCTCCAGTCC GTCTGGCCTC GCTGTTGCTG
2501 GAGAGGCCAG GTCCTGCCTC AGTGCTTTTG GCTTGGCTGT TTCGTTTGCC
2551 ACGGATGTCT TTCTTTCCCC AGATATCAAC ATGGCTTGCT GGTCAATCGC
2601 TTCAGGTCTT CAAGTCTTGG GTCAAATGGT GGCTTCTCAG TGAAGTCTTA
2651 TTTGACCACA CTAAAAATTG CACCATCTCA CCCCATTGT CCTTTTCTTG
2701 CTCGATTTTG TTTTACCCC ATAGCACTTA ACACCTTACA ACAAGCTATA
2751 TATTTTGCTT ATTTCAGTCA TTCATTTAAT AACTATTCGC ACCTATTTGT
2801 GTGCCAGGCT ATGTGTGCCC CCACTGCATG GGGGCAAACA TCTCTGCCCT
2851 TGTGGAGCTT CCATTCTAAG GGGGGAGATA ATAAACACAT TTATAAGTAA
2901 GAGAGTATGT CAGATAAGTG TATCATCTCC TGTCACAGTG AGTTAAAATC
2951 TGGTGTTTAA TCTCCATGAT TAGACTGAGC TTCCTAAAC TGGAGTGGA
3001 GCTGATTTTC ACCTCCTTGT CCCTGATATC TTGAGGGAGA TCAGGATCTC
3051 TCAGGCCCTT CCTGCTCAAA ACATAGGACA CACTTGACTT TTCTGATATC
3101 CTTTCAGCGC CAGTGGGGAG ATGTTGAAGT TCAAATATGG AGCGCGGAAT
3151 CCTTTGGATG CTGGTGCTGC TGAACCCATT GCCAGCCGGG CCTCCAGGCT
3201 GAATCTGTTT TTCCAGGTAA CAGCCTACCC TGCCAACCTT GCTCACCTGT
3251 GTGTGTCCTT GGAATCTCCT TGTCACCTAC CTTTGCTTTT ATTTATTTGT
3301 TTATTTATTT AGAGTCTCAG TCTCTCAGGC TGGAGTACAG TGGTGCAATC
3351 TCAGCTCACT GCAACCTCCG CCTCCTGGGT TCAAGCGATT CTCCTGCCTC
3401 AGCCTCCAGA GTAGCTGGGA CTACAGCCGC CTGCCACCAC ACCCGGCTAA
3451 ATTTTGTATT TTTCTTTTTA GTAGAGACGG GGTTTCACCA TGTTGGCCAG
3501 GCTAGGGTCG AACTCCTGAC CTCAAGTGAT CCACCTGCCT TGGCCTCCTA
3551 AAGTGCTGGG ATTACAGGCA TGAACCGTGC CCAGCTTGCT TTTATTATAG
3601 GACCAGGGAT AATATTTTAG GGGAAATTCT GTTTTGTTTT GTTTGAAACA
3651 AGGTCTTCTG TCGACTCTAG GCCTGTGCCA CCATGCCTGG CTAATTTTTT
3701 AATTTTTTGT AGGGATGGGG TCTCACTGTG TTGCCCAGGC TGATATAGAA
3751 CACCTGACTT CAAGTGAGCC TCTTGCTTG GCCTCCCAA GCCTGGGGT
3801 TATAGGTGTG AGCCACTGCA CCTGGCCCTC TATTTAGAGT TTTATATGCA
3851 CTGATTCTTT TGGAAAAAAG ACACTGTGCA GAAGTAGATA GCTGAACCTG
3901 CCTTAGAAGG GAGATCTTTT CATATTTCTC ACACCTTACA CTTCTGTACT
3951 AAAGTTTATT CATTCATTGA TTGATTGGTT GCTTGCAAGA CAGGGTCTTG
4001 CTCTGTGGCT CAGGCTGGAG TGCATTGGCA CAATCACGGC TTACTGCAGC
4051 CTTGACCTCC TGGGCTCAAA CGATCCTCCC ACTTCAGCTT CCTGAGTAGC
4101 TGGGACCACA GGTGTGTGCC ACCATACCTG GCTAATTTTT GTATTTTTTG
4151 TAGAGATGAG GTTTCACCAT GTTGCCCAGG CAGGTCTCGA ATTCCTGGGC
4201 TCAAGTGATC TACTTGTCAC AGCTTCTGCA AGTGTTGGGC TTACAGGCAT
4251 AAGCCCCTGT ACCAGGGCAA GTTTGTCCTT TTATTGAAGA AAGAAAAATA
4301 AATGAACAAA GATGCTTTTT AAAACTACAA TTTCTGTGGG TATAATCCTA
4351 TTCATTTTCA TTGCAGGGAT GTTTATTTTT TAAGATTTTT TTTTTTTTTT
4401 TTTGAGACAG AGTCTTCGCT GTCGCCCAGG CTGGAGTGCA GTGGCGCGAT
4451 CTCGGCTCAC TGCAGGCTCT GCCCCCGGG GTTCACGCCA TTCTCCTGCC
4501 TCAGCCTCCC ACGTAGCTGG GACTACAGGC GCCCGTCACC TCGCCCGGCT
4551 AATTTTTTGT ATTTT TAGTA GAGACGGGGT TTTACTGTGT TAGCCAGGAT
```

FIG. 3B

```

4601 GGTATTTTTT AAGATTTTAA AAAAAGTTTT GATGAATACC ACACCTGTTT
4651 AACCCTCATT CCTCTCAAGA TACACATTTT TGTCACCCCA GATGCGTTAA
4701 AACTTAATAT CATAAGATTA CTTCCAAATA GATTTTTAAT TCTTTTGTTT
4751 CTGATGTATG TGAACACTG GTGAAGTAGA AATCCTTGTT TGATTTATGT
4801 ATTCGTAAGT CAGGGGGACA ATAGAGACCA TGAAGATTTA GAATTGAATC
4851 CCAGTCCCAG CACTAGTTAG CTGCATTACT TTGGGTGAGT CAGTTACCTT
4901 TTCTGAGTCC ATTTGCTATT CTTTAAAATA GGTGTAGACC TGTAATGCCA
4951 GTATTTTCGG AGGCTGAGGC GGGCGGATTA CTTGAGGTCA CGGGTTTCGAG
5001 ACCAGCCTGG ACAACGTGGT GAAACCCTGT CTCTACTAAA AATATAGAAA
5051 ATTAGCTGGG CATGGTGGTC GCATGTACCT GTAATCCCAG CTACTTGAAA
5101 AGCTGAAGCA GGAGAATCAT TTGAACCCGG GAGGCGGAGG TTGTCGTGAG
5151 CCGAGATGGT GCACTGCACT CCAGCCTGGG CGACAGAGTG GGTAAGACTC
5201 CATCTCAAAA CAAAACAAAA CAAAAGAAAA CAAAAAAAT AACATAGAGG
5251 TTGTAGTACC TAATCCACAG GGTGTGTGTG AGGATTAGAT GAGATATTCTG
5301 ATTTAAAGCA CTTAGCACCT TGCCTGGCTC TTAGTAAACT CTTATAAAAA
5351 AATGGTAATT ATTGTTAATA CTCAGCATAG AATAGTATTA GTTATAATAT
5401 TAATACTAAA TTTGTTTCCT TAATAGTAAT TATATTTGGG AAGGTAGTTA
5451 TGTAGGATAC CTGTAAGATG ATGAATGATG AAGTATTCTT GATAACTTTT
5501 TTTTTTTTTT CAAAATATTG GTATTGGGTG TTAAACAGA TGAGAGTGGA
5551 AACAAATTGA AAGCTTAGGT TTTTCTGTGG GACCATCCCC ATCAGCATTT
5601 TAAGTCTTGA CATATCTTTC ACAAATGAAT AGTCTGTCTT TAACCTTAGA
5651 TGGCTGGAGT GCTGCCACGT TTCAGCCCCT TTATCATGCT ACTTTAAAAAT
5701 ATCTCCAAC TGTGCGCGT GGTGGCTCAC GCCTGTAATC CTAGCAATTT
5751 GGGAGGCTGA GGTGGGTGGA TTGCTTGAGG TCAGGAGTTC GAGAGCAGCC
5801 CGGGCAACAT GGTGAGCCCC TCCGTTTCTA CTAAAACAC AAAAAATAGC
5851 TGAAGTGTGAT GGTGTGTGCC TGAGTCCCA GCTACTCGGG AGGCTGAGGC
5901 AGGAGGATCA CTTGAGCCCT AGAGGCAGAG GTTGCAGTGA GCTAAGATTG
5951 TGCCACTGCA CTTGAGCACT TCAGCCTAGG CGACAGAGCA AGACCCTGTA
6001 AATTAAAAAA AAAAAAAAAA AGAAAAGGAA AAAAATTTCC AACTTATTAA
6051 GGGCTTATAG TGTGCTGATT ATGTAATAGT TATGGCTTCC AATGTGTCTG
6101 GCATAGAACT GGCATGTTTC TGAGTATCTC ACTTCAGCCT CATGACAGAG
6151 GTAAGGACTA TTTTAAATTT AAACTTTAAA TAGGAGGCAA CAGGCCAGGT
6201 GTGGTGGCTC ACACCTGTAA TCCCAGTACT TTGGGAGGCT GAGGCAGGTG
6251 GATTGCTTGA GTCCAAGAGT TCAAGACTAG CCTGGGCAA ATGGTGAAAC
6301 CCCATCTCTA CAAAAAATAT AAATAATTAG TCAGGCATGG CCGTGTGTGC
6351 CTGTAGTCCC AGCTACTCAG GAGGCTGAGG TGGGGGCATC TCTGGGGCCC
6401 CGGAGGCAGA GGTGTAGTG AGTTGAGATT GCAACACTGC ACTCCAGCCT
6451 GGGCAACAGA ACGAGACCCT GTTTCTAAAT AAATACATAA ATAGGAGGCA
6501 ACAGATATAG ACAGATATGG AGGTAGGTAA GGCCTTGCCC AAGATCATAC
6551 ACGTTGGGTT TTGCAGATGA GGCCAAGATC AGACTCCATC TTTGGTTGGT
6601 CTGACTCCAA AGGCTGACCA CATAGCCATT GGGCCACAGC ACCTGTGCAC
6651 GTCAGAATTT ATTAAGTATA TCTTGATTTT AGTCATTATA ACAGGAAGAC
6701 TTATGGGTAA ACCCTCAGTT CATCTCTTTT TAATGCTGAG ATCCCCCTGC
6751 CCAGTAAAGC TATTATTGCA AGTATAGTAT ATACCTATCA TTTGCCTTGA
6801 GTTATCAGGT AAGGATGCTG TTTGTTCTTT TCCCATATAG TGCTGTTTGA
6851 ATGAGGTTGA GATACAGTAG CAATTTTGTT TTCCATTGAG GTGAGTACCT

```

FIG. 3C

```
6901 TAGACTGAGT GTCATTTTGT CTTTTTTACT TCTACTCAAC AGGATTTTCCT
6951 GACATGTTTCG AGGTCAGTGA TTGTCAGACT TTCTGAGCCA GCAAAATTTTC
7001 CCAAATTGCT GGGTAGACAC AGGTTTTTCCA ACTTTTTTATT TTGCCAAGTA
7051 AGGATATATA AAAAAAAAAAT AAAAAGAAAG ACCTATTATT TTCTGGCCCT
7101 TGTATTTTCAT AAAGGGCATT TTAAGAAACA ACAAGACAGG AAGAACATCA
7151 TCTCAGAATA AAGGACCATT TTTAAATTTG AATACATTTA GTTTTATAAA
7201 AAAGATATCA TGTGGTGTTC ATTTTTTCTC ATTTCACTGC AGGCTGTTGA
7251 AAACTTTGTT AAGAACCAGT ACTATATTTG GGAACCCCTG CTTTAATTGA
7301 TCTAAACTCT TGAAGAATAG AAGAAACAAA GCATTTTATT TTTCTGAGTT
7351 ACTGGCAACT ATTACTAAAG TGACAGATAT GGTGGCCTTG AATGCAGTGC
7401 TTCCCAAACC TGATTGAGGT CTGACTCTCT TGGGGACCAG GGTCTCATTC
7451 TGTTGCCCAG GCTGGAGTGT GGCAGCACAA TCTTGGCTCA CTGCAGCCTT
7501 TACTTCTTGG GCTCAAGTGA TCCTTCTACC TCAGTCTCAC AAGTGGCTAG
7551 GACTACAGGA CCATGGCACT ACACCTGGCT AATTTTTTTT TGTTTGTTTG
7601 TAGAGATGGG ATCTCGCTGT GTTGCCCTGG CTGGTCTTGA ACTCCTGGGC
7651 TCAAGTGATC CTCCACCTT GGCCTCCCAA AGTGCTAGTA TTCCAGGTGT
7701 GAGCCACCTC TCCCTGCTGG GGAAC TTGTT AATAAAACAG ATTCTAGGCT
7751 ACAGTCTGGA AAATTCTAAT TCATTTGGTT GTGGGGGAGG GGGGCATAGG
7801 ACCAGAGAAT GTGTTTGT TTGTTGTTTGT TTTTCTTAAA TTCTCCAGTG
7851 CTGTTGTGAT TCAAATGCAG CCGTCTGT TCTGTTATCA AGTGCTGTGT
7901 AACAAAGCAC TCACAAAGTT TAAAGCAACA ATGATTTATT TTTTCTTAGG
7951 ATTCTGTGGG TTGGCTGGAC TCAGCTAGGT AGTTCTGCTT CATCCTGTGA
8001 TGTCAGCTGG GGTCACTTGT GGGGCTACAT TCAGCTGGGA TTATGTCTGG
8051 GACTGGAACA TGTGGGTGCT GACTGCTGGC TGGGGCACCT TAGTGTTTCT
8101 CACATGGCCT CTCTTCTCCA TGAGGTCTTT CAGTAGTATA GCCCAGGACT
8151 CGTAACTTTT TTTTTTTTTT TAAGACAGAC TGTCGCCCTG TCGCCAGGC
8201 TGGAGTGCAG TGGCACGATC TCTGCTCACT GCAACCTCCG CCTCCTGGGT
8251 TCAAGCAATT CTCCTGCCCC AGCCTCCCGA GTAGCTGGGA TTACAGGCAC
8301 GTGCCCTCCAC GCCCGGCTAA TGTTTGCATT TTTAGTAGAG ATGGGGTTTC
8351 ACCACGTTGG TCAGGCTGGT CTCGAACTTC TGACCTCGCG ATCCGCCTGC
8401 CTCGGCCTCC CAAAGTGTG GAATTACAGG TGTGAGCCAC TGCACCTGGC
8451 CGACTCGTAA CTTTTTTTGT AAGTAATAAA TATTTTAGGC TTTGTGGGTC
8501 CTGTAGTCTC TGTTGCAACC ACTCAACTTG GCCATGGTAG CACAAAAGCA
8551 GCTAAAGACA ATATGTAAAT GATGGGTGTA GCTGTGTTCC AGTAAACTT
8601 ATAAAAAGTC CGTGGGCTGG ATTTGGTCCA AGGGCTACAG ATTGCACACC
8651 CCTGGTCTAG CCAAGCATC TGTGCATGGT GGCTGGCTTC CCAAAGTGG
8701 AAGCTGCTAA GCTGCCTTTT TTTTTTTTTT TTTTTTTTTT GAGAGGGAGT
8751 CTCCTGTGT TGCCTAGGCT GGAGTGCAGT GGTGTGATCT CGGCTCACTG
8801 CAACCTCCAT CTCCGGGTG CAGGCAATTC TCATGCCCTCA ACCTCCAGG
8851 TAGCTGGGAT TACGGGTGCC TACCACCACG CCTGGCTAAT TTTTGTATTT
8901 TGGTAGAGAC AGGGTTTCAC CATGTTGGCC AGGCTGGTCT CAAACTCCTG
8951 ACCTCAAGTG ATCCACCCGT CTTGGCCTCC CAAAGTGCTG GGATTACAGA
9001 TGTGAGCCAC CGTGTCTGGC CGCTTGACAA GCTTCTTAAA GGCAGTGGC
9051 TGAAGTGGCA CAGTGTCACT TGTGTCACAT TCTTTTGGTT GAAGAGAGTC
9101 TCAGAGATGG CACAGATTCA AAGGCAGGAG AAATAGACTC CAGCGCTTAA
9151 AGTAAGGAGT AGCATGTGCC TACAGAATTG GAGGAAGTGT TGGAGGCCAT
```

FIG. 3D

```
9201 CTTTGAAGAG AGACCACCAC TATCCATGGC TTGGCACGTG GGAATCACTG
9251 CTCTATACCA GGGTTGCAGA CTCATGTCTT TGGGGGCCAG GCAGTGAGTA
9301 TAAATGAGTC AAGTGGGCCA GTTGGAAGAT GGAGTCAGAC CTGCAGTGAA
9351 CTCCCAAACA CATCTGCTAC CGGGAGGGGC AGCATTACTC AGCTCCAGCT
9401 CAGCGTCATC AGGCAGGAAG GCGAGGCAGT GTTGCCGGAT GTGCCAGTGT
9451 TTCAAAAGAA GCCAGAGACT CCATTTTAT TTTTGTGTAT GGAATCTCCT
9501 GATTTTGAAA TATTGGCAGA TAATTCAAAT TATCTTAAAC ACTACAGGCC
9551 AAACAAAACA TATCTGTGGG CTAGAGACAG TCTGCCAGTT TGTAACTATT
9601 TCTCCAGATC ATGAGTAAAT TTGGCTTTAC GATGGTCACT CAGTTCCTAT
9651 TACTCTAGGT TGTTCAAATG AATTAAAAAA GCTGAAATTA TATGAATAAA
9701 CCCCTGGGCA CACATGAAAG AAGTGAAAAA CCCATTGTTT CCTATTGTAG
9751 AAACATGGAA GCATGTCAGA GCCAGAGGAT CCAGAGGAAA TATTCTCACT
9801 AGCCTCAGAC CCTCAGGAGT GAGGGAGCTT TTCTTGTTAA TGGCCACGCT
9851 TGTGCAGTTT TCCTTCCCAG GTGCTGGTGA AAGAAACCCA CAGTCTTGGA
9901 ATCATGGAAG TGATACCATA ATGACTGTCA GTTGACGTTG CTTTAAAGAA
9951 TGAAGCCACA GAATTGTGCT GTTAGCATGT CGTGAGCAGT TAGTTGAGTT
10001 GGTGGCTTGT AATTTACTCT GTGTGGATGT TATTGATCAA AGCTTTTCAT
10051 TATTGACAGT GTCTCCATCT GCTGTTTGCT GTTTTtaggg GAAACCACCC
10101 TTTATGACTC AACAGCAGAT GTCTCCTCTT TCCCGAGAAG GGATATTAGA
10151 TGCCCTCTTT GTTCTCTTTG AAGAATGCAG TCAGCCTGCT CTGATGAAGA
10201 TTAAGCACGT GAGCAACTTT GTCCGGAAGT GTAAGTTTGG GGAACTTTTT
10251 CTTGAAACT GTCTGAGAG AGAAAACTA GAAAGATGCT TGAGGCAGAA
10301 TGAGTTACTG GTTGATAGTA GTCGGTAAGA ACTCTGGTTC TATATAAGAC
10351 AGATCCAGGT TCAAATTCAG GCTGCACCTC TTATAGCTGG GAGACCAGGT
10401 AAGTTGGGCT TCTTGTTGTC AAGCGACAAA CTTAATTCAA AGACTGAATT
10451 TAGGCCAGGT GCAATGGCTC ATACCTATAA TCTCAGCCCT TTGGGAAGCT
10501 GAGGTGGGTG AATCGCTTGA GCCCAGGAGT TCAAGACCAG CTTGGGCAAC
10551 ATGGTGAAAC CCCATCTCTA CAAAAAATAC AAAAATTAGC TGGGTATGGT
10601 GGCTTGCAAC CGTGGTCCCA GCTGCTGAGG AGGCTGAGGT GGGAGGATCA
10651 CTGGAGCCCG GGAGGTTGAG GCTCAATGAG CTGTGATTGT GCCATTGCAC
10701 TCCAGTCTGG GTGACAGAGT GAGACCCTGT GTGAATAAAA GAGTGAATTT
10751 ATTGGCTCAT GAAACTGAGA AATCCAGGAA TGAGTTAAGT TTTAGCTTTA
10801 GGCATAGCTA GTTCCAGAGA CCTCAATAAT ATCCCGTGGC CCTGTCCTTA
10851 TACTCACTCA GGGCTGACTT TCTATTAGGC AGAGTAGGCA CGGTGCTTAG
10901 GATCTGTGAT ATTTAATTTT AATGAATTTA ATTACTTTTA ATTAAGTAA
10951 TTAAATTTTA ATTTGTTTTT AAATTATAGG AAAAATGAAT ATAATAATGT
11001 ATAATGATTC TGGATTACAT TCATCTTTAT ACTAATGTAG TCATAAAATA
11051 TAATTTTTGT TTTTTTTGGA GACAGAGTCT TGCCCTATTA CCCAGGCTGG
11101 ATTGCAGTGG TATATCATGG CTCACTGCAG TTCAACCTT CTAGGCTCAA
11151 GCAATCCTTC CACCCAGTG GCTGGGACTA CAGGCTCACA CTACCACGCC
11201 CAGCTAATTT TTGCTTTTTT CTCTGTAGAG ATAGGGTCTT ACTATGTTAC
11251 CCAGGCTGGT TTCAAACCTC AGGCTTGAAG CAGTCTTCCT GCCTCAGCCT
11301 CCCAAAGCTT TGGGATTACA GGTGTGAGCC ACCATGCCTG GCCCCATAAA
11351 ATATAATTTT TGAATTCCTT TTTGTTTTTA ATGGAGGAAG GGGCTGAGGA
11401 AGGCAAAAGT ACCTAGGGCC TATGAAGTCA TATATTGGCC TTGCCTTCAC
11451 CCTGTTTCTG ACTTTGCTTG ACTTCCATGT GATGAGGCAG TTGGCTGTTA
```

FIG. 3E

```

11501 GTGTCCCAGT TTCATACTCT TACATTAGTG TTTTCAACC AGTGGGTGAT
11551 TTGACGTTTT CGGTTGTCAG AGCTAGTTGG GGGTGGTGGT GTGTGAGTTT
11601 GGGGGGAAGG GTCCTACTGT CAGTTAATGG GTGAGGCCAG AGATGCCACC
11651 AAACACCTTA CAGTGCACAA AGCAGCCCC ATAACACAGA ATTATGTAGC
11701 CCACAATGCC AACAGTGCTG AATTGAGAA ACCCCACCTT GTACAACATT
11751 GCTGTGCAAC CAACCACCCT AAATATTACT GACTTAAAC AATAGTCACT
11801 GTGGCTGGGC GCGGTGGCTC ATGCGTGTA GCCCAGCGCT TTGGGAGGCT
11851 GAGGCGGCGG ATCACTTGAG GTCAGGAGTT CCAGACCAGC CTGGCCAACA
11901 TGGTGAAACC TTGTCTCTAC TAAAAATACA AGAATTAGCT GAATGTGGCA
11951 GCGGGCGCCT GTAATCCAG CCATTTGGGA GGCAGAGGCA GGAGAATCGC
12001 TTGAACCTGG GAGGTGGAGG TTGCAGTGAG CCAAGATCTC ACCATTGCAC
12051 TCCAGCTTGG GCAATGAGTG AGACTCTGTC TAAAAAATAA AAAAAAGTTA
12101 TTGTATTACC TCTTGTGTGT GTAGGTTAAT TGGACTCAGC TGGGGATTCC
12151 TCTGCTCTGT ATTACATTGG CCAGGATTGC AGTCACCTGG GGCTCTCCTG
12201 GGCTGGAATG TGTGAGAGGG CTTACTCAGT GTTTGGTGCC CTGGCTTGGA
12251 GGCTGGGCCC AGCTGGGCCT CTCTCTCTTC ATGAAGTTTC AGGGCCTTTT
12301 GCTGTCCACA TGGCACCTCT ATGTGGTCTC CAAATCAGAA GTCAAGGAAC
12351 TACAGCCTGT GATGCCTATT TTGTAAAGAA GGTTTTACTG GAACACAGCC
12401 CTACCCATGT GTTTGTACAG TGCCTATGGC TGCTTTCACA TCATAACAGC
12451 ATTTTATTTT ATTTTATTTA TTTTTTTTTG AGACAAAGTC TACTCTGGC
12501 TGGAGTGCAG CAGCACAATC ATAGCTCACT GCAGCCTCCA ACTCTGGGC
12551 TCAAGCAATC CTCCTGTCTC AGCCTCCTCA GTAGCTAGTA CTACAGGCCC
12601 ATGCCACCAC TAATGGCTAA TTTTTTAATT TTGTGTAGAG ATGGGACCTT
12651 GTGAGATTGC CTAGGCTGGT CTTGAACTCC TGGCCTCAAG AAATCCTCCC
12701 ACCTTGGCCT CCCAAAATGC TTGGATTACA GGCATGAGCC ACTGTGCCCC
12751 GCCACAACA GCATTTGAGT AGTTGTGATA GAGACCAAT GGCCTACAAA
12801 GCCAAAATA GTTCCTGTTT GGCCCATTTT GAAAAGGCTT GCTGACCTCT
12851 GAGCTACATG GTCTCTCTAG CAGGACAGCC TCGACGGTAG CTCAGGTTTC
12901 CAAAACACAA AAGTGAAGC TGCCAGGCTT TCTTAGGGGT TATCCTAGGA
12951 GGGACATAGG ATCTCTTTGA CTGCATTTTA TTGTTTGATG CATGCTCTGG
13001 GGCTGCTCAA ATTCCACCTG AGAGGAACT ACACAAGGTC ATGAATCCCA
13051 AGAGGACTGG GGCATTGGGT GCTATTTTGG GAGACTGGCT ACCACACCCT
13101 GCCCAATGGT AATCTTCCCT TATCTAGATT AATACAACCC CAGGGAAGAT
13151 TCTAACTTGG CTCTGCTTTG GGTCAATTTG CTCCCTGGAG GTGAGGTGTT
13201 GTGATCGGTT TTGTTGGAAT GCCCAAAGGG GTCAGGGCAG TGTGATTACC
13251 AGGACCTCAT GGAATGGGGG ATGCGTGGTT ATGCAAAGGA GCCGGGGATG
13301 CTGGGTAGAA AAAAAATCAG CATATGTTCA CTATAGTGCT CTTCAGTATT
13351 TTACATGTAC TTTGTTCTCA GTTTTCTCAT CTGTAAAATA GGAATAATGT
13401 ATATCCTTTT TTTTTTTTTT TTTTGGAGT CTTGCTCTGT TGTCCAGGCT
13451 GGAGTACAGT GGCACAATCT CAGCTCACTG CAACCTCCGC ATCCCGGGTT
13501 CAAGTGATTC TCCTGCCTCA GCCTCCTCAG TAGCTGGGAC TACAGGCGTG
13551 CACCACCACA CTCAGCTAGT TTTTGTATTT TTAGTAGAGA TGGGGTTTCG
13601 CCATGTTGGC CAGGCTGGTC TCAAACCTCT GACCTCAAGT GATCTGCCTG
13651 CCTCGGCCTC CGAAAGTGCT GGAATTACAG GCATGAGCCA CCACGCCCAT
13701 TGGGAATAAT GTATATCTAA TGAGGCTGTG TTGGAATTGA ATGAGTTAAT
13751 GCACAGACCA GATTTGTCAT GTTGCCTGGC CCATAGGAGA CAATAAATGG

```

FIG. 3F

13801	TACCCAGTAT	TAATAACTGT	GAATGTCAAC	AACATTTAAT	ATATTGTATA
13851	TCTTCAAAAT	GTACTTGAGG	TATTTGTTCA	TCATTCTGTT	TTTGTTTGAA
13901	TAAGCTCGTG	CCTTCTTTTT	GTGAATATTT	AAATTTATAA	GTAGCGAGTG
13951	GGAGGGGAAG	GAAGTTATGT	GATGAGGCTA	GCTTACTGAG	CCATCTGCAG
14001	GCACCTTCAT	TAGTCTTGAG	ACTGTCCTCT	GGTACTTAA	CAGCAGTGAA
14051	TTATCTAGAA	TCATTTAGTG	ATCAGAAGAC	TTGGTTTAGT	GGAATGTAGA
14101	TTTTTTTCTA	ATAGACCCCT	CTTCCAGGGA	AATGTTTCAT	ATTTTTGAAG
14151	AGGTTTCCTG	GGGAGTGTTT	AAGAGGCCAT	GATTGAAAAT	GGGTGATTAC
14201	ATTAGTGTGT	TTTCTATTCC	TCCCCTTTTT	GAGTTTCTGT	TTTGGAATGT
14251	AAGCTTTGTT	TTTCTACGTG	GAGAAGGGTC	CCTCAGCTGC	TTCTGCCCAG
14301	GTTTTTTGAA	TCTTCCTATA	GGGATGGAGA	TTTTCTTTGG	GGACTGTTAG
14351	AGAAAATGGA	ATAGAGTGTA	GCTCTGAAGG	AGAAGGATGT	CTCCAGCAGA
14401	AGTACCTCTA	GCCTTGGGCC	AAGGGAGGGA	AGGGAAGGGA	ACGAGCATCT
14451	GGGAACCAGG	GAAGGGATTT	TTGTCTTTCT	TAATTACTCT	TACATCCCCA
14501	GTGCCCAAAA	TAGTGTCTGG	CATATGTTAA	GTCCTTAGTA	AATACTTGTT
14551	GAATGAGTGT	ATGCTCAGTG	AACAAAATAA	ATGGCAAACA	TTAAGCACAG
14601	TATCAGATAA	TTTGTGTAAA	AAATATACAG	CAGTGTTATA	CTAAAACCTG
14651	CACAGAGGCC	AGGTGCAGTG	GCTCACGCCT	GTAATCCCAG	CACTGGGAGG
14701	CCGAGGTGGG	CAGATCTTTG	AGCTCAGGAG	TTTGAGACCA	ACCTGGGCAA
14751	CATGCTGAAA	CCCTGTCTAT	ACAAAAAATA	CAAAAAGTAG	CTGGGGCATG
14801	GGGACGCACA	TCTGTGGTCC	CAGCTACTTG	GGAGGCTGAG	GCTGGAGAAT
14851	TGCTTGAAGC	TGGGAGGTGG	AGGTTGCAGT	AAGCCAAGAT	TGTGCCACTG
14901	CACCCCAGCC	TGGGTGACAG	AGTAAGACCC	TGTCTCAAAA	CACAAAACAA
14951	CACCCCCTTC	AAAAAAAATC	CAAAACCACC	ACCACAACAA	AAAAACTTAC
15001	ACAGAAAAGT	GTTGATAATT	GTCAAAATTG	GGCTGTTATT	GGCAATTTGA
15051	CAGTAGCTGA	ATTACTACCA	TTTGAGCTAT	ATTCACTATA	GATAAGATCT
15101	TCAATATATT	TACAACCTTA	GTACTAATGG	GAAAATGATA	ACTTTTGAAA
15151	AGTTTTTTTT	TTTTCTTATT	GCAAACAATA	CACAATACAA	TGTTAAATAT
15201	AGAAGGTTAA	ACGTGCATCT	GAGTCTGTTT	GGGCTGCGAT	AATAGATACC
15251	TTAGACTTGG	CAATTTATAA	ACAATAGAAA	TTCATTGCTG	ACAGTTGTGA
15301	AGACTGGGAA	GTCCAAGATC	AAGGCGCCAG	CGAATCTGGT	ATCTGGTGAT
15351	GGCTCCCTGC	TTCAAAAATG	GCGCCTTCTT	GCTGCATCTT	CACCTGGCAG
15401	AAGGGGCAAA	CATGAGTCCT	TCAGCTTCTT	TTTTTTTTTT	TTTCTATGTT
15451	TAAAACTTTT	GGTCCGGCGT	GGTGGCTCAT	GCCTGTAATC	CTAGCACTTT
15501	GGGAGGCCGA	GGCAGGTGCA	TCATGAGGTC	AAGAGATCGA	GACCATCCTG
15551	GCCAACATGG	TGAAACCCCC	CCGTCTCTAT	ACTAAAAATA	CAAAAATTAG
15601	CCAGGCATGG	TGGCGTGTGC	TTGTAGTCCC	AGCTACTCAG	GAGGCTGAGG
15651	CAGGAGAATT	GCTTGAACCT	GGGAGGCAGA	GGTTGCAGTG	AGCCAAGATT
15701	GCGCCACTGC	ACTCCAGCCT	GGCAACAGAG	TAAGACTCCG	TCTCAAAACA
15751	AACAAACAAA	AAAAACAAAA	AAAAACTTTT	ATTTTAGGTT	CATGGGTAAA
15801	TGTACAGGTT	TGTTATGTAG	GTAAACTTGT	CTTGGGGTTT	GTTATAGATT
15851	ATTTCGTCAC	CCAGGTACTA	AGCCTAGTAA	CCAATAGTTA	TTTTTTCAGA
15901	TTGTCTCCCT	CCTCCCACCC	TCTGTCTCTT	AGTAGGCTCC	AATGTCTGTT
15951	GTTCCCTTCT	TAGTGTCTCT	GTGTTCTCAT	CCTTTAGCTC	CCATTTATAT
16001	GTGAGAACAT	GTGGTATTTG	GTTTTCTGTT	CCTGCATTAG	TTTGCTAAGG
16051	ATAATGTCAG	CCTCTTTTTT	TTTTTTTTTT	TTTTTTTTGA	TACAGAGTCT

FIG. 3G

```
16101 CGCTCTGTTG CCCAGGTTGG AGTGCAGTGG TGCGATCTTG GCTCACTGCA
16151 ACCTCTGCCT CCCGGGTTCA AGTGATTCTC TTGCCTTAGC CTCCTGAGTA
16201 GCTGGGACTA CAGGTGCGCA CCACCATGCC AGGCTAATTT TTGTATTTTA
16251 GTAGAGATAG GGTTTCACCA TGCTGGCCAC GCTGGTCTCC AACTCTTGAC
16301 CTTGTGATCC GCCGGCCTCG TCTTTTTCCC AAAGTGCTGA GATTACAGGT
16351 GTGAGTCACT GCACCCGGCC CAATGTCAGC CTCTTTTTTTA GGAAGTGAT
16401 TTAATCACTT CCCTAAAAGT CCTACCTCGT TTTTTTTTTT GGTTTTTTCT
16451 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TAGGTAGAGT CTTGCTCTGT
16501 CACCCAGGCT GGAGTGCAGT GGTGCGATCT TGGCTCACTG CAACCTCCAC
16551 CTCCTGAGTT CAAGCAATTC TCCTGCCTCA GCCTCCTGAG TAGCTGGGAT
16601 TATAGGTGCC TGCCACCACG CCTGGCTAAT TTTTTGTAT TTTTAGTAGA
16651 GTTGGGGTTT CACCATGTTG GCCAGGCTGG TCTTGAAGTC CTGACCTCAA
16701 GTGATCTGCC CAAAATGCTG GGATTACAGG CGGGAGCCAC TGTGGCCAGC
16751 CCCTGCAAGT CCTACCTCTT AATAGTATTA CACTGGGGAT TACATTTCAA
16801 CATGAATTTT GTAGGGGCGA GGGGCACAAA CGTTTAGAAT ATAGCACATC
16851 ACATACATAG TGAGAGAAAA ATCCCTCAAA ATCTTACCTG AGACAATCAC
16901 TGCCAACAGA TTGCTGTATA GTGTGCCAAT TTTGTTTGTG TGTGTGTGTG
16951 CCTTAAAAAT ATTTATTATG GAAATTTAAA AACGTACCCC AAGGTGGCCA
17001 GGTGTAGTGG CTCACGCCTG TAATCCTGGC ACTTTGGGAG CCCGAGGTGG
17051 GTGTATTACT TGAGGTCAGG AGTTTGAGAC CAGCCTGGCC AAAATGGTGA
17101 TACCAGTCTC CTAAAAATAC AAAAATTAGC CGGGTGTGGT GGGCACCTGT
17151 AGTTCCAGCT ACTCGGGAGA CCAAGTCATG AGAATTGCTT GAACCTGGA
17201 GGCAGAGGTT GCAGTGAGCC AAGACCATGC CACTGCACTC CAGCCAGGGT
17251 GACAGAGTGA GACTCCATCC TAGAAACAAA CAAACAAACA AACAAACCAA
17301 CTAACCAACC AGAGAAAACCT CCCTGTCTGT AAGGAGTATG TGTTCTAATG
17351 GATACTGAGC CATCTTGTTT TGTTTAACAT GTGCCTAATG TTCTTTTATA
17401 TGGGCGGACT TGTAGGTTGT TTCAACTTTT CTGTTGATGA ACCTTTAGGT
17451 GGTTTCTGAT TATTTTTGTG TTACAACAGT TTTCATCATT CACATCTTTG
17501 TATGCATCTT TTTTGAGCAC ATGTGCAAGT ATTTCTGTGG ACAATGGATG
17551 ATTCCTAGAA ATTGAAAGTT TGGATTACTG TGTTCCAAAA AAGGAAGCAA
17601 TACACCCAGC TATGTTGGCT TTTGCTCTTG GGTCCAGATG ATTATCTGAC
17651 AAAGTTATTC TCTGATTGCA TTTTCTTTTC TTTTCTTTTC TTTTTTTTTT
17701 TTGAGATGGA GTTTCGCTCT TGTTGCCAG GTTGGAGTGC AATGGCGCGA
17751 TCTCGGCTCA CTGCAACCTC TGCCTCCAG GTTCAAGCGA TTCTCCTGCC
17801 TCAGCCTCCT AAGTAGCTGG CATTGCAGGC ATGCGCCACG ACACCTGGCT
17851 AATTTTTTGT ATTTTTAGTA GAGATGGGAT TTCTCCATAT TGGTCAGGCT
17901 GGTCTTGAAC TCTTGACCTC AGGTGATCCA CCCGCTTCAG CCTCCCAAAG
17951 TGCTGGGATT ACAGGCGTGA GCCACAGTGC CTGGCCCTCT GACTGCATTT
18001 TCACAGTGTT TTGGGTCCTT ATCTCTACCT CAGTACCTCA ATATTCACTG
18051 CCCACTGGGC CCTTAGATAC TGCAGCTAAA AGTGCACAGG GGTGGAGTGA
18101 TGTGACGGTT TTGGGGTCAC AGAAGCAGCT GGTATAGAGA GAAGTTGTGA
18151 AGTTTTTTTT TTTTTTCCTG AGACAGAGTC TCGCTGTATC CCCTAGGCTG
18201 GAGTGCAGTG GCTTGATCTC GGCTCACTGC AACCTCTGTC TCCCTGGTTC
18251 AAGTGATTCT TATGCCTCAG CCTCCCGAGT AGCTGGGATT ATAGGCATGT
18301 GTCACCATAC CCAGCTAATT TTTGTGTTTT TAGTAGAGAT GGGGTTTCAC
18351 CATGTTGGCC AGGCTGGTCT TGAGCTCCTG ACCTCAGGTG ATCCGCCAC
```

FIG. 3H


```

18401 CTGGGCCTCC CAAAGTGCTG GGATTACAGG CCTGAGCCAT TGCGCCTGGT
18451 CTTTTTTTTT TTTTTTTAAG TAATCATAGG CTTGAATGTA GCCTCTCATC
18501 TGTTACACCTT AATAATCCAA AAGCCTTTAG ATAAAGAAAT GGAGATTTGG
18551 AATGGCTTCT CAGAATTCCA AGAGAGTATT GTCATGGTTT TGCCTGCAAA
18601 GCACCGTGGT CTGTCTCCTT GTGCAGTTGA GAAAGCTGGT GGTGCGCCACT
18651 GACAGGCCCA GAGTTATTAA GTTGGACACT GCTTTAAGCA ACTTTGTAAA
18701 CAATCCAAGG CATACTAGAG AATTAGGAGA GATTGGCTTT GTGTATGAGC
18751 AATAACAAAA TCAAGTTCAA TCCAGCAAGT TTTTGGGGAA TTATAATTCA
18801 AAACCTCAAAT ACTTGATCTG GAAGAACTT GGAAAGAGGG AAGGAAGACA
18851 GGCTTGTTAC AGCATTGTCA GGGTAAAAGG AAAATACCGT GCAGCTTTTA
18901 ATTTTGCTTC TTCATGGCAT TCCCACATGTA GGTGCCCTAG ATTTGTTTTT
18951 TACAGTGGTC ACGACTTCAT GTGGATCCAC CCACCACTCT TGCTTGTTTC
19001 CCCAAGGGAC CAAGGGAAGG TGTATTGAGG ATGATTGCTG AAGTGAGGGG
19051 TGGGGTCTGT GGCTGAGAAG ACTCTCAATA CCGCGGCACT CATTATAAGC
19101 CTCTGACACA GGAGATTTCA ACTCCACCCG TGCAACAAAG GAACAGGGTG
19151 GGCAAGAGTA GTTACAGTTG CAGGCTGAGT GCGATGGTTC ATGCCTGTAA
19201 TCCAGTGCTT TTGGGAAGCC AAGGTGGGAG GATTGCTTGA GTCTAGGAGT
19251 TTGAGACCAG CCTGGGTGAC ATAATGAGAC CCTACCTGTA CAAAAAATT
19301 TTAAAAATTA GCCAGATTGG TGGTGTGCGC CTATAGTCCC AGCTACTCTG
19351 GAGAATGAGG TGGGTGAGGG TCCCTTGAGT CCAGGAGTTC GAGGCTGCAG
19401 TGAGTTATGA TTCTATGATT TCACCACTGC ATTCCAGCCT GGGCGACAGA
19451 GCAAGATTGT GTTCTTTTTT TTTTTTGAGA CGGAGTCTCA CTCTGTCACC
19501 CAGGCTGAAG TGCAGTGGTA CGATCTCTGC TCACTACAAC CTGCACCTCC
19551 CAGGTTCAAG TGATTCTCTC CCTCAGCCTC CCGAGCAGCT GAGATTAAAA
19601 GCGGCCGCTT GTGTGCAGCT AATTTTTGTA TTGTTAGTAG AGATGGGGTT
19651 TCATCATGTT GGTCAGGCTT GTCTTGAACCT CCTGACCTCA GGTGATCCAC
19701 CCGCCTCGCC CTCCCAAAT GCTGGGATTA CAGGCGTGAG CTACTGCGCC
19751 CAGCCATTTG TGTCTCTTAA AAAAAAACT AAGAAAATGA AAAAAATGAC
19801 ATTGGCCAAT TCATTAAAAAT GCCACTCACT GACTGTGGTA TGAAATGGCT
19851 TTCCCTTTGA TGGACCGAGT CTGTCTCATT GTGTGAGCCA CTTGCAGGGC
19901 TGAGTATGAC TCTGGAATGT AGCTCCTAAC CTTATCTGCT GCCCAGCCAT
19951 TGAAATGGCC ATCCCTTCCA GTTCCAGAA GATTCCAGTG TGTGTTTGGG
20001 ATTTTAAGAC AGTCTCTTGG TCTTCAGTGT GGCATCTTTC TGCCGGATTT
20051 TCCAGGATAA TTTTGATTAT AAGCATTGCA TTGCCCTTGG TGTGTAATGC
20101 CTGTGTATGA TGCTGTTCCC TTGTAACGTG CAGGATTAAA TTTTGGGTC
20151 AGCCACTGCT GCTCCCCCTC ATTCCTGCAG GTCATTAGAG TCATCGTACA
20201 TTTAGCGATG TCTCAGATCA GTGTATCTAG AGTGTTAATA AACATGTTAG
20251 ATTCCAAATC TACTGTCCAT TTAATCCATA CTTCATACTG TGAGGATCTC
20301 TGAAGTAAAG ATTAGACTTG GAAAAATAAT AAGACTGTAT GGTAAAGAAA
20351 CTATAGTTGC AAATCCATTT GGACATGTAG TATGTCAGCC CTGCAGAGCA
20401 GATGTCAGAA CCCCATTTAG TTCTCTGAGT GCTAAGCCCT TCTGCCCACC
20451 ACGCTGTTTT TTTTTTTTGA GATGGAGTCT CGCTCTGTCA CTCAGGCTGG
20501 AGTGAGTGGT TGTGATCTCG GCTCACTGCA AGCTCTGTCT CCCAGGTTCA
20551 CGCCATTCTC CTGCCTCAGC CTCCCAAGTA GCTGGGACTA CAGGTGCTCA
20601 CCACCATGCC CAGCTAATTT TTTGTATGTT TTTGGTAGAG ACGGGGTTTC
20651 ACTGTGTTAG CCAGGATGGT CTGGATCTCC TGACCTTGTT ATCCACCCGC

```

FIG. 3I

20701	TTCGGCCTCC	CAAAGTGCTG	GGATTACAGG	CGTGAGCCAC	TGCTCCTGGC
20751	CCCCACGCCT	TTTTTTTTTT	TTGGAGACAG	AGTTTCACTC	TGTCACCCAG
20801	ATTGGAGTGC	TGTGGCACAA	TCTCAGCTCA	TTGTGTCCTC	TGCCTCCCAG
20851	GTTCAAGTGA	TTCTTGTGCC	TCAGCCTCCT	GAGTAGGTGG	AATTACAGGC
20901	GTGCACCACA	ACACCTGGCT	AATTTTTTGT	TTTTTAGTAG	AGATGGGGTT
20951	TCACCATGTT	GGCCAGGCTG	GTCTCGATCT	CCTGACCTCC	AGTGATCCAC
21001	TTGCCTAGGC	CTCCCAAAGT	GTTGGGATTA	CAGGCGTCAG	CCACCATGCC
21051	TGGACCCCTC	TGCCCCCTTA	AGCACTGCCA	CATATTAGAT	CTACGAAGGC
21101	TTTATGGATA	CAATCCAAGG	AAGATGAACC	TTGGGCTAGT	GGGATAAAAC
21151	TAAGCGCATG	TAGTTAGAAT	GGAATGATCT	GGAAACCAGG	TCCCAAGTTG
21201	GTCTAAATTA	GACTCATGTT	GACTATGTCA	CACTGTAAAC	CAGTCTAAAT
21251	GCTAATAAGC	ATGCTTGACC	AAACACTGCC	CTGCAGCCTT	CAGAGAGGAA
21301	GAAGGAAAAC	ATAATTTGTA	TCCTCTCTCC	CTATTTTCTG	AGTCTATGGG
21351	ATTCAAATTG	TAGCTGCCAT	GGAAACTGTA	CTTTGGAATT	TCTAGAGCCC
21401	TTAATTTTAA	CTTAACATAT	AAAAACACTT	TTGTACTGAT	TTTATAATTA
21451	TTCATGATGG	ATGAGAAAGT	GAATGTCTTT	GACAGTGAGG	GAAGCTATCC
21501	GAATGCTATT	TTCTTTTTTT	TTTTTCTTTC	ATAAAGATGC	ATATATTTGC
21551	ATGCTTTTATT	TACCTGGGGC	TAACCTCTGC	ATCTTTTGCA	GATTCCGACA
21601	CCATAGCTGA	GTTACAGGAG	CTCCAGCCTT	CGGCAAAGGA	CTTCGAAGTC
21651	AGAAGTCTTG	TAGGTTGTGG	TCACTTTGCT	GAAGTGCAGG	TGGTAAGAGA
21701	GAAAGCAACC	GGGGACATCT	ATGCTATGAA	AGTGATGAAG	AAGAAGGCTT
21751	TATTGGCCCA	GGAGCAGGTA	GGAGGATTTT	AACATCATGC	TTTTCCACTT
21801	TCTGTACCGG	AGTGTTCAAT	GCAAAGACGA	TAATCTGCTG	CACTGGCGTC
21851	TAGGATCAAG	CACGTTTTTC	TCTGTGACTC	TATATTTAAT	TATAGTTGGG
21901	GCAAAAAGGT	CTCTCATGTT	CTTAGCTCAT	CTTCTTGAAC	TGATGTTGGC
21951	TAATTTTGAA	GGCTCACAAA	TTCTCTTGA	TGTATCATGT	TTCTATCGTT
22001	GTAATTTTATT	TCAGAACCAA	GGTGGCCTTT	TAGCTAATGA	ATTTAAGATG
22051	ATCTTTTATG	ACCATTAGCT	GAGGACTCAG	GATATACATA	TGGTGGGGTG
22101	AATCAGATTG	CTTTTGTACA	CGCTTTAGGT	ATTTGTGTTG	TGGGCATATG
22151	GATTTGGTTT	TAAAACAGGC	CTTTGAAGAA	ATCAAATAAC	ATTCTTTGTT
22201	ATGTGGCTAG	GGAGTTGCTT	GTTTGAGAGC	AGGTAGAACG	TTATCTTTTTT
22251	TGTTGTGGTA	TTTTTCTTTC	TTTTAAACAA	GGCTACTGTC	TCTAGACATA
22301	TTGATTCATT	TGCTGTGTTT	TAGAGAGATG	GCCGTCAGCC	TTGGAATTCA
22351	GAGAGTAATT	TATTACTTAC	AGACATTTTA	GTGCACATGA	TATGTCTGAT
22401	AATGTACCCA	GCTCTGCAGG	AAGCTTGCAA	AAGGAATAGA	AGTCCCATGG
22451	TTGCTATTTT	CAGTGTTTAA	AAACAACCTT	GGAAAGTGGA	GGAAAAATGC
22501	AAATGTATAA	AGCAGGTGCT	TACCAGCTAA	AGTATCACAG	AAGTGGGAGA
22551	GCAATTAGCA	AATTAATTAA	CGATGATGTG	AGGGGAGATG	TTGTGGGTGA
22601	GCAAGGGACA	GTTAGGGACA	GTTCTCACCG	ATGGGGGGAA	ATGTAGGTTT
22651	TCGGCAGAGA	GAAGTGATGA	GAACATGTTG	GGTAGAAGTG	TGACATTCTG
22701	GAGTACTAGA	ATGCTATGCA	AGTGTGTGTG	TGTGGGTGTG	TGTGTGTGTT
22751	CAGTGGTTCA	GAACAGACTG	GGAAATGGCG	AAATGAGGAC	ATTTGGGTGG
22801	GGAGGGGGAA	ATGGGTGGGA	AACTCAAGAA	CCTTTTTTTA	AAAAATTGTG
22851	GTAAATATAA	TATAACATAA	AGTGTACCAT	TTTAACCATT	TTTAAATGTG
22901	CAACTGAGTG	GTATTCAGTG	CATTCATGAT	GTTGTACAAC	CATGACCGCT
22951	CTCCATTCTT	AGAATTTTTT	TATCATCCCA	AACAGAAACT	CTCTATCCAT

FIG. 3J

23001 TATACAATAC CTCCCCATTC CCCCAGAAGC CAGTTTTTTGA ATTGCAGTTT
23051 ACTTTGTGAG GCTGTTGGGG ATTATTTAGG CCTCTGGAAG GAGGAGGTTG
23101 GGATCAGAGT CTGGCCCTGT GGACTTCAAT GACTTTGTGT GCCTCCAAT
23151 CAGAGAAGCA GCGGAGGGCA GGAAGCTGCT TGTCAGAATC TGAGAGTGAT
23201 GTGGCTTCTT TGTTTAGCAA TAAAATGTGA GCACATAATA GAAAGGAAAA
23251 GTGACAGGAC ATGGCAGATA ATTTGGAAGA GAGGAGTGGA AGATGCTCAC
23301 TCAGCCTCCC AGCTCCTGAG AAAGAACTGT GTCTCATCAG TTCATACTAC
23351 CTGAGCATCT GTTGTATCTG GTGTGTTTCT AGGTCCTGGA GAAGAGGCAT
23401 TACGTGTAGC CCTGACCTTG TGATGCTTAT GTTTTTGATG GGAAATAGTG
23451 CGTGTAAGAA GAAAATAATC CAACAGGCCA CACGGCAGGC AAACAATAGA
23501 GATATTCAAA TAGGTATACC TTCTCCAGG TGAATGGCCT GAAATGACCG
23551 TGTGGAAGTG TGGGCTGGGG GCTTATAAAA TTATACACAT ACAGGCGCTA
23601 ACTAAAGCCG CCTATTCAAT CCTTAAGAGG ATGCATAGAA AAGAAAAGTA
23651 GGGTCCTTAA CTGAGCCATT TGGAATTTAA GGGCATGAGA GAAGCCAGCA
23701 CAAGCAGTGA AGGGAAGGAA AAGAAGTGCC CGAGAGGAGG GAGGGATGCT
23751 GTTCTGCAGA CAAGGCCTGC CGCCTGGGAG AGGCCCCGAC GCCACCCAG
23801 GGTCTCTGTA CAGCTGGAAG GGTCTTCAG AGACTGTTTA TATTTTATTT
23851 ATTTATTTAT TTATTTATTT TGAGACAGAG TCTCTGTCAC CCAGGCTGGA
23901 GTGCAGTGGT GCGATCTCAG CTCACTGCAA GCTCCGCCTC CCAGGTTTAC
23951 ACCATTCTCC TATCTCAGCC TCCCGAGTAG CTGGGACTAC AGGCGCCTGC
24001 CACAATGCCC GGCTAATTTT TTTGTAATTT TAGTAGAGAC GGGGTTTTAC
24051 CTCGTTAGCC AGGATGGTCT TGATCTCCTG ACCTCATGAT TCGCCCACCT
24101 CGGCCTCCCA AAGTGCTGGG ATTACAGGTG TGAGCCACTG TGCCTGGCCG
24151 ACTGTTTCTA CTATTTTAGA GAGAGGGTCT CACTGTCATC TGTGCTGGAA
24201 TGCAGTGATG CAGTCATAGC TCACTGCACC CTCAAACCTC TGGGCTTAAG
24251 CGACCCTCCC GCCTCAGCCT CTTAAGTAGC TGGGACCATA GGCATGTGCT
24301 GCCACACCCA GTTAACTTTA TTATTTATTT ATTTATTTAG AGAATGAGTC
24351 TCATTCTGTT GCCCAGGCTA GAGGTGCAGT GGCACGATCT CGGCTCACTG
24401 CAACCCCGCC TCCCAGGTTT AAGCGATTCT TCTTGCTCAG CCTCCTGAAT
24451 AGCTGGGATT ACAGGCACCT GCCACCACAC CTGGCTAATT TTTGTATTTT
24501 TAGTGCAGAG GGGGGGTTTC ACCATGTTGG TCAGGCTGGT CTCGAACTCC
24551 TGACCTTGTG ATCTGCCTGC CTCGGCCTCC CAAAGTGCTG GGATTACAGG
24601 CGTGAGCCAC CGTGCCCGGC CCACTTTATT ATTTTAAAAA CATTGTTTTA
24651 TTTTATTTTT TTTGAGACAG AGTCCGCTGG AGTTCAGTGG CCGGATCTCA
24701 CTCACTGCAA CCTCTGCCTC CTGGGTCAA GTGATTCTTG TGCTTCAGCC
24751 TCTCTAGTAG CTGGGACTAC AGGCGGGTGC CACCATGCCT GGCTAATGTT
24801 TTTTGTATCT TTTTAGTAGA GACGGGGTTT TGCCATGTTG GCCAGGCTGG
24851 TCTCGAACTC CTGACCTCAA GTGATCTGCC CACTTTAGCC TCTCAAAGTA
24901 CTGGGATTAC AGGCGTGAGC CACTGTGGCT AGCCCCCAGC TAACTTTAAA
24951 AAAAAATTTT GTGGGCCGGG TGCAGTGGCT CACGCCTGTA ATCCCAGCAC
25001 TTTGGAGGCC AAGCAGGGCG GATCACTTGA GGTCGGGAGT TTGAGACCAG
25051 CCTGACCAAC ATGGAGAAAC CCTGTCTCTA CTAAAAATAC AAAAAATTAG
25101 CCGGGTGTGG TGGTGCATGC CTGTAATCCC AGCTACTTGG GAGCTGAGGC
25151 AGGAGAATTG CTTGAATCTG GGAGGCAGAG GTTGCAGTGA GCTTAGATCA
25201 CGCCACTGCA CTACAGCCTG GGCAACAAGA GCGAACACTC CGTCTCAAAA
25251 AAAAAAATA AATTATGTAG AGGTGGGATC TCCCTATGTT GCCCGGACTG

FIG. 3K

```
25301 GTCTTGAACCT CCTGGCCTCA AGTGATCCTT CCATCTCCCC CTCCCAAAGT
25351 GTTGGGATTA CAGGCATGAG CCACCCCTCC TGGCTGAGAC TGCTTATTTT
25401 ATTTATTTTT AATTTTTTTT GTTTTGAGAC TGCTTATTTT AATGGAAGCT
25451 TCAGGGGTCA GACGGGTCA GACAGAGTCA TTGGTGAGCA AGCAAAGGTG
25501 TAGACTGTTC AGTTCAGCCT TCCTTGAGCA CCTTTTATGT GCCAGACAAA
25551 AGAAGGATCA GCATATCAGG TGCAGTAAAT TATTGGGGTT ATGTTGGTGT
25601 TTCCCAAATG TGTTAGATTT ATCCCTGGTA GTGTTAAATC TCATGATTTT
25651 AGGTAGTATA TGGACAACCT ATGTAAAAAC ATTTAATAGT TTAATATTAA
25701 CTAGCATATC AAAACCTGTG ACTTTGCTCA CGCCTGTAAT CCCAGCACTT
25751 TGGGAGGCCA AGGCGGGAGG ATGGTTTGGG CCCAGGAGTT TGAGGCCAGC
25801 CTAGGTAACA TGGTGAGACC CTGTCTCTAA AACAAAACAA AACAAAACAA
25851 ACAAACAAAC AAATAAACAA ATCCCCTGTA ACTTGTTCTA ACAATAACCT
25901 AAACAATTTT TTATTTAAAA TTAAATAAAA AAATTGAAAC AGTAACCATT
25951 TTTTTTTTTT TTTTGGAGA CAGAGTCTTG CTTTGTCCACC TAGTCTAGAG
26001 TGCAGTGGCA CAATCTCTGC TCACTGCAAC CTCTGCCTTC AAACAATTCT
26051 CCTGCCTCAG GCTTCTGAGT AGGTGGGATT GATTACAGGT GCACTCCACC
26101 ATGCCCAGCT AATTTTTGTA TTTTATAGTAG AGACGGGGTT TCACCATGTT
26151 GGCTAGGCTA GTCTTGAACCT CCTGACCTGC AGTAGTCCAC GTGCCTTGGC
26201 CTCCCAAAGT GCTGGGATTA CAATCACAAA TTTATAGAAA AGTTGCAAGT
26251 ACCATGTAGT CAGGGTTCTT AAGAGAAATG GAACCAGTAG GAGATAGATA
26301 TATAATCATC TCCTAGGATT ATAAGTTGAC ACATAAGACT AACCGTCACA
26351 TACAGTATAA ACAACTTTTT TTCTTAAACC ATTTGATAGA TACACACACA
26401 CTGATATACA TAGAATATAT ATACACACAC ACAGAATGTA TATACACATA
26451 GAATATATGT GCATACAGAA TATATACACA GAAATATATA TGTACACATG
26501 CATAGAATAT ATTTACATAT ATATGCATAT ATATAATTTA TTTATTTTAA
26551 GCAGTTGATT TATACAGTTT TTGTTTTTGT TTTTTTTTTG AGACAGAGTC
26601 TCACTCTGTC ACCCAGGCTA GAGTGCACTG GCGAGATCTC AGCTCACTGC
26651 AACCTCTGCC CCCGGGTTC AGTGATTCTC CTGCCTCAGC TCCACAAGTA
26701 GCACACCACC ATGCCCAGCT AATTTTTGTA TTTTTTTTAG TAGAGACGAG
26751 GTTTCATCAT GTTGGCCAGG CTGGTCTCGA ACTCCTGACC TCAAGTGATC
26801 CGCCCGCCTT GGCCCTCCAA AGTGCTGGGA TTTCAGGCGT GAGCCACCAC
26851 ACCTGGCTCC CATAATGTCT TTTAGAATAA AACGATCGAG TTGAGGATCA
26901 CACGTGACAC TTAATTGTCC TGTCTCTTTA GTCTCCTTCA ATCTGGAGCA
26951 GTTCTTTGAT TTTTCCTGGA CTCTCATGAC CTTGACAATT CTGATGATTA
27001 TAGGCCAGTT ATTTTGTA AAA ATTTGAATTT GTCTGATGTT GCTTATGTTT
27051 AGATTTAGGG TCTTGGTCTT TGGCCGGAAT ATCTCAGACA AGATGCTCTG
27101 TTCTTATTGC ATCAGAGCAG AAGACTCTCT GTTTCAGTTG ATCACATTTA
27151 TGTTGATGCT CACTTTGATC ACTTGATTAA GGTGGTGTCA GTTATGCCTT
27201 TCTACTTGTA GGGTTACTCC TTCTCCTTC GTGATTTTAT TTATTTTATT
27251 TTTCTTAGAG ACAGGGTCTT GCTTGGTTGC CCAAGCTGGA GTGCAGTGGT
27301 GGGATCTTGG CTCACTGCAG CCTTGAACCTC CTGGGCTCAA GTAATCCACC
27351 TGCCACAGCC TCCTGAGTAA CTGGGACTGT AAGCGAACAC CACCACACCC
27401 AGCTACTTTT TGTATTGTAG AGATGGGGTC TCACTGTGTT GTCCAGGCTG
27451 GTCTGTAACCT CCTGGCCTCA AGCAGTCTTC CGGCCTTGGC CTCCCGAAGT
27501 GCTGGGATTA CAGGCATGAG CCACTGCACC CAGCCTCCTT TGTAATTAAA
27551 AAAGTATTTT ATGGGGAGTT ACTTTCAAGT GATGGAAATA TTTTATATCT
```

FIG. 3L

```

27601 ATGTGGACTT GGATTTTCCT ATTCAGTCA GTGAGTTATA ATCCATTTCT
27651 GTCACTAGTT TTATACTTAA ATTGTTCCCA ACTTGGCCAC TGAGAACCTT
27701 TTTAGGTTAG CTTTTGTGTC CTTTTCACAT GTCTCCAAGA TTCATTGAAT
27751 ACTTTCCTGC TTTCTGGTAT AGCAAGATGT TCAGGTTCTT TTGGTACTTT
27801 TACTTTCTCT GCCCTGGCTC TGGCATCAGT CATTTCTCAG AGGAGCCCTG
27851 TGCCTTTCAG TGGACAATGG TGTTTAGAGG CCAAGATCTG GACATTGGGT
27901 GTTTTCATTG CTACCGGTGT GTCACTACTC CCAGACCCCT TTCAGTGGAC
27951 AGCATAAGG AATACACATA CGTATATACA ATATATCCAC CTACACATGT
28001 GCGTGCACCT ACACACACAC ATATACATTA CATCTATATT TGTGTATCCA
28051 TGTCTATATA TTGAAAATTG TGGCTGGGCA CAGTGGCTTA TGCCTTTAAT
28101 CTCAGCATTT TGGGAGGCTG AGGCAAGAGG ATCACCTGAA GCCAGGAGTT
28151 CAACACCAGC TTGGGAAACA GAGAGAGACT CTGTCTCTAC AAAAATAAAA
28201 AGGGAAAACC ATGAGTTCAC ACCCGTGCCC CCAGTTCCAA TCCAACCTCA
28251 CAGGGTTCAT TTTAGTTTTT ACCCTTTCCA TGTGTGTAAT TCTCTTCTCT
28301 GACATTATAC CCTTAATATG TTTACTTATT TTATGCATCT GTATGCATCC
28351 AATCTACTGT CTTTGTGGT ATCCCACCTC CCCTTGGTGG GTCCAGATAA
28401 TCTGCTCTGG GTTGCCCTTT CACGTGGATG TCTTCCTTAC CCTGTGTGGG
28451 CCTGTGATAC TGGGCTGCCC CCACACATGA GTGCTGCCCT CCTCACGTTG
28501 CTGGGACGG CACTGTGTCC TGGGCCACCA TGACTTTTCT CATAACTAGC
28551 GTGGATGCTT ACCTTGTTCC ACACCAGTGA ATGGCTTCAG GAAGAGAAGA
28601 GGAAGAGAAA AATATTTACA TTAAAGAAA GGAGTTTAA AGAAATATGT
28651 TAGGTAAAGA ATTGAGCAGG TAATATACGG AGCTGGCAA AATTGTGACC
28701 AAAGTAGGTG AATGATTGAG ATTTATGCAA TTCTGGGCTA AGTGACAGCC
28751 CCTTCCCTTT CCCTTCCCTT CCCCTTCCCT TCCCTTTTCT TCCCTTTCCC
28801 TTCCCTTTCC TTCCCTTTCC CTCCCCTTC CCTTCCCTTT CCTTCCCTTT
28851 CCCTCTTCTT CCTTCCTTCC TTCTGTTTTT TTTTCCCTTC TTTCCTTTGC
28901 CTTTTTTTTT TTTTAAAGC TAGAAACATC AGTTTAGGCA TAAAGACAGA
28951 GGAAAAGGCT TCTTTTCTCT CTCACAGTTC TTTATAATTG TCTAAGCAGT
29001 TTCTTTTTTT CCTAGTTTTT ATTTTTTGAG GAAGAGCGGA ACATATTATC
29051 TCGAAGCACA AGCCCGTGGA TCCCCAATT ACAGTATGCC TTTCAGGACA
29101 AAAATCACCT TTATCTGGTG AGTCTTTACA TCTGTCTCTC TGGAATTAGC
29151 CTAGCACTCT GACACTCAGA TGCCTGTGGT AGAACTGAAT GTTGTCTTG
29201 CCCATGTGGT CTCATTTCATG CAAAGACTTT CTTACCTTAC AGGTGTCTCC
29251 CTGTTTTTCT CGTTATAAAG ATCAAGAGCT AACCATTATA GAAACAGCCT
29301 CATTGGGCTG AACGTGGTGG CTCACGCCTG TAATCCCAGC ATTTTGGGAG
29351 GCCGAGGCGG GTGGATCACG AGGTCAGGAG ATCAAGACCA TCCTGGCTAA
29401 CACAGTGAAA CCCCCTCTCT ACTAAAATA CAGAAAAATT AGCCGGGCAT
29451 GGTGTCGGGT GCCTGTAGTC CCAGCTACTC AGGTGGCTAA GGCAGGACAA
29501 TCGCTTGAAC CTGGGAAGCG GAGCTTGCAG TGAGCCGAGA TTGCGCCACT
29551 GCACTCCAGC CTGGGTGACA GAGCAAGACT CTATCTCAAA AAAAAAAAAA
29601 AGAAAAAAAA AGAAACAGCC TCATTGACAG TTGGATATTG TAGCTGTGGC
29651 TTTCAGGCAA TAATAGGGAA TCATTTATTG GGGAATAGTC TGTCATTATG
29701 TATAAGATAA TCTTGCTTTA ATTTTTAAAA ACTTCCTGTG TTAGCTTGCT
29751 TAGGATTAAA AAAATGATAA TAGTGCATGG TTGTTATAAG AAAATGCAAA
29801 CACTGCAGAC ATGCATGAAG TTGAAGGGAA AGCCCCCAT TTTCTTTTCC
29851 TTTTCTTTTT TTTTGAGACA GAGTCTCGCT TTGTCACCCA GGCTGGAGTG

```

FIG. 3M

29901	CGGTGGCACT	ATCTCGGCTC	ACTGCAATCT	CCACCTCCCA	GGTTCAAGAG
29951	ATTCTTCTGC	CTCAGCTTCC	CTAGTAGCTG	GGATTACAGG	CACGTGTCAC
30001	CACGCCCAAC	TAATTTTTGT	ATTTTTAGTA	GAGATGGGGT	TTTACCACGT
30051	TGGCCGGGCT	GGCCGCAAAC	TCCTGACCTC	AAATGATCCA	CCTGCCTCGG
30101	CCTCCCAAAG	TGTTGTGATT	ACAGGAGTGA	GCCACTGTGC	CCGGCCTCTC
30151	CGTTTTATTT	TCTAATCCTC	CTCCCTAGGG	GAAGAAATGT	TAAATGGTTA
30201	CATAAGCTTT	CCCTTTCTGA	CCCTTAACTG	TGCTCTGTAG	GAGCATGGTG
30251	GGGGATGTTT	CTTTTCTTTT	CTTCTTTTTT	TGAGACCAGG	TCTCACTTTG
30301	CCACCCAGGC	TGGAGTTCAG	TGGCATGAAC	ATGGCTCACT	GCAGCCTCGA
30351	CTTCCTGGGC	TCCAGCAAAC	CTCCCACCTC	AGCCTCCCGG	GCATACACCA
30401	CTGTGCCTGG	CTAATTTTTG	TATTTTTAGT	AGAGACGGGG	TTTTGCCATG
30451	TTGCCCAGGC	TGGTTTCGAA	GTCTTGAGCT	CAAGAGATCT	TCCTGCCTTG
30501	GCCTTCCAAA	GTGCTGGGAT	TACAGGTGTG	AGCCACCATG	CCCAGCTCCG
30551	GTGGGGGATA	TTTCTATATC	CACATGTGTA	TAGTTTACTT	TATAAAAATG
30601	GTATGTTACT	CTGTGCTTGG	CTCTCCAGCT	TGCTGTTGCC	TTTCACCAGT
30651	GTATCCCAGA	CATCCTTTCT	TCCTTGTCAG	TAACGCAGGT	CTACTTTATT
30701	CTTTGAGCAG	TGGCATAATT	TTCCCTGATG	TGTATATATC	ATAAGTTAGA
30751	GAATGCTAAA	ATTCATTTTG	GGGCCTTGTT	TAGGTTCTTG	AGGGATTAAA
30801	TTCTTAAATT	TAACAAGTGT	ATCCTGGAAA	CAATTTTTGT	TCCTGATTCA
30851	GCCCTTAAAA	GAGGACTATC	ATGTTACCTT	GAATGGAGAT	AAACAGGCTC
30901	ACGTAAGAGA	AAAGGGTAAG	AGGGATGAAC	TCCCATTAT	CTTAACTTC
30951	TACTGGCCCG	TTTTTGGGGA	ATTTGCTGCT	TTTATTCCTG	ACCTAAAATA
31001	AATAAGTTTA	TGTGTCTTGG	TTTCATATTA	GTTGAGAACC	CAGTGCCTGG
31051	AGAGAAGTTT	TCCTTGTCCT	CTGAGTGAGG	ACATTCACAT	ATGAATCTAT
31101	TGGCAGACTG	GCTTTGACTG	ACCACACGTG	CCTTCAGAAC	CAATGCCACA
31151	GCTCTTAGGT	TTATGGCCTG	AAACACCCTT	TCCTTACATA	TTGCCTTAGA
31201	AACTTTCCTT	CCTTGAGACA	TGGGGCATGG	AACCCTCACC	TTACAGATG
31251	ACCTTGGTGT	GTTTCTAGGG	TTGCTGGTGT	TCCAGGACAT	CTGTTGCAGA
31301	TGCAGTATTT	ACCTTGTGCT	CTCTGCATCA	TAAGCAGCTT	CTCATGTTTG
31351	AATGTATTAA	CAGACTTTTA	ATTTTTTTTA	TTTTTTGAGAC	AAAGTCTCAC
31401	TCTGTCACCC	AGGCTAGTGT	TACCCAGGCT	GGAGTGCAAT	GGCTCAATCT
31451	CAGCTCACTG	CAACCTCCAC	CTCCTGGGTT	CAAGCGATTC	TCTTGCTCA
31501	GCCTCCCGAG	TAGCTGGGAT	TACAGGTGCA	TGACACCACG	CCCTGCTAAT
31551	TTTTGTATTT	TTAGTAGAGA	CGGGGTTTCG	CCATGTTGGT	GGGGCTGGTC
31601	TCAAACCTCCT	GACCTCAGAT	GATCTGCCCC	CCTTGGCCTC	CCAAAGTGCT
31651	GGGATTACAG	GCGTGAGCCA	CTGCGCCTTT	TCTTTTCATT	TTTTTTCTGA
31701	GATGGAGTCT	TTCTCTGTCA	CCAGGCTGGA	GTACAGTCAT	GCAATCTCAG
31751	CTCACTGCAA	CTTCCACCTC	CTGGGTAA	GTGATTCTCC	TGTCTTAGCC
31801	TCCTGTGTAG	CTGGGACTAC	AGGCGTGTGC	CACCGTGCCC	AGCTAATTTT
31851	TATATTTTTA	GTAGAGACGG	GGTTTTGCCA	TGTGGGTTAG	GCTGGTCTTG
31901	AACTCCTGAC	CTCAGGTGAT	CCACCCGTCT	TGGCCTCCCA	AAGTGCTGGG
31951	GTTATAGGCG	TGAGCCACTG	TGCCCAGCCT	CAGGCTTCTT	TATTAAGAAG
32001	AAGTTCGGGC	CAGGTGTGGT	GGCTTACACC	TGTAATCCCA	GCAATTTGGG
32051	AGGCCGAGGT	GGGCAGATCA	GGAGGTCAGG	AGATCGAGAC	CATCCTGGCT
32101	AACATGGTGA	AACCTCGTCT	CTACTAAAAA	TATAAAAAAT	TAGGCAGGTA
32151	TGGTGGCGGG	TGCCTGTAGT	CCCAGCTACT	CGGGAGGCTG	AGGGAGGAGA

FIG. 3N

```

32201 ACGGTGTGAA CCTGGGAGGC GGAGCTTGCA GTGAGCCCAG ATTGTGCCAG
32251 TGCAGTCCAG CCTGGGTGAC AGAGCGAGGC TCCGTCTCAA GAAAAAAAAA
32301 AAAGACGTTT CCTTGAAACA ACAGGGCTTT TGTTTGTTTT GTTTGTGTTT
32351 TGTTTGTTAT TGTGTTTTTA GATACGTATT TTTTCTTTC TTTTTTTTTT
32401 TTAAGTGATG ATGTCTCTGT TGCAGTGGCA TGATCATAGC TCACTGTAAC
32451 CTCAAATTGC AGGGCTCAAG TGATTCTCCT GCTTCACCTT CCTGATTAGC
32501 TGGGACAACA GGTACAAACC ACCATGCCTA GCGAATTTTT AAATTTTTTCA
32551 TAGAGACTAG GGTCTCACTA TGTGCTTAG GCTGGTTTCG AACTCCTGGC
32601 CCCAAGTCAT CCTCCTGCCT TGGCTTCCCA AATTGTTGGG ATCACAGGCA
32651 TGAATCACCA CACCCAGCCT ATTTTATAGT ATTTTAATTC GAGCTCTACA
32701 GGAGGTTTAG AACACTAGCT TGTGAAGATA AACTTCATTT TCAAGGCCAC
32751 ACAGAATCTA AGTGGTCCTG GAATTAGGAA GGGCTTTGAT TTTTGGACC
32801 AAAGTTGAGA GTCCACAGTT TTCTGGTCTA CCTTGCACTG CTCCATAAAC
32851 TCATATTTCT TTTCTCTGAG CTGAAGAGCT CCCCTTCTTG GTGTCTAGTC
32901 TCAGGCAACT TATTCTTAAA AGTAAGCATT ATTGAAATGC TTTGGGATTT
32951 TCACATCATC AAGGTCCATT TTGGTAGAGG CACTGACAGA TTTTGAGTGT
33001 TCTGTGTGAA GGAAGTCAGT TGAGGATTTA GTGGTCCATG TGGCAGGCTA
33051 CTGCTCAGTA GCTTCAGGGA AACCAGTCTG TGCCTCCCCT GTGGCCAGTG
33101 AGGATGATCA GAGGAGTCCC AGCAGGAATG CCCAAATGTA GTTTTCTTAC
33151 ATGTTGATGG GAGTGCATTG TTTCATGTCT AAACAGTTCT CAAATCACAT
33201 CTTCAGGAGG GTACTATCTG GGCACTTTGA TAATTTCTCA CTTTGATGTC
33251 ACCGTTCTTA TTACCATCAC CTAGTTTTGT CATAGTAGAA ATAACTTTCC
33301 TTTTCTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTTTT
33351 GAGATGGAGT CTTGCCGTGT TGCCCAGGCT GTAGTGCAGT GCGGTGTTCT
33401 CGGCTCACTG CAACCTCTGC CTCCCGGGTT CTCCTGCCTC AGCCTCCCGA
33451 GTAGTTGGGA TTACAGGCGT GTGACACCAC GCCCGGCTCA TTTTGTATT
33501 TTCAGTAGAG ATGGGGTTTC ACCACTTTGG CCAGGCTGGT CTTGAACTCC
33551 TGACCTTG TG ATCCGCCCAC CTTGACCTCC CAAAGTGCTG GGATTGCAGG
33601 TGTGAGCCAC CACGCCTGGC TTTTTTTTTT TTTTTTTTGA GACAGAGTCT
33651 TGCTCTGTTG CCCAGGCTGG AGTGCAGTGG CGGGATCTTG GCTCACTGCA
33701 GCCTCCACCT CCTAGGTTCA AGCAATTCTT CTGCCTCAGC CTCCTGAGTA
33751 GCTGGGATTA CAGGTGCCCC CCACCATGTC CGGCAAATTT TTGTATTTTT
33801 AGTAGAGACA GGGTTTCACC ATGTTGGCCA GGCTGGTTTC TAACTCCTGA
33851 CCCAGGTGA TCCGCCTGCC TCAGCCTCCC AGAGTGATGG AATTACAGGC
33901 ATGAGCCACT GCGCCTGGCC ACCTTTGTCT TCTTAGTTGT GGATTTAACT
33951 GCTGTGGACA TCTGCTTGGG CATAGCCTTC CCGGAGTACC TCTTGATTG
34001 GGACTGTCTG TGGGTTTCTG TGCTAGGACA GGCTCCCAGA TGTAGGAGGC
34051 TTCCCAATG ATCTCACCAC TGGCATCGGC ATCCTTAGCT TCTACTCAGC
34101 TTTTCCATCT GCCATCTTGC AAGATGGAAG GTTGTTTTGT TTTTGTTTTT
34151 GTTTTTTGGT TTATTTTTTT TGAGATAGAG TCTCGCTCTG TTGCCAAGGC
34201 TGGAGTTCAG TGGCGCAATC TCGGCTCAGT GCAACCTCCA CCTCCTGGGT
34251 TCAAGTGATT CACCTGCCTC AGCCTCTGGA GTAGCTGGGA TTACAGGCGC
34301 GTGCCACCAT GTTCGTTTAA TTTTTTGTAT TTTTAGTAGA GACGGGGTTT
34351 CACCGTGTTA GCCAGGATGG TCTCGATCTT CTGACCTCAT GATCCGCCTG
34401 CTTCAGCCTC CCAGAGTGCT GGGATTACAG GCGTGAGCCA CCGTGCCAG
34451 CCTAGGAGGG TTCTTAATGC AGCTGTTTTT TGGAGTCTG GTTGCCCTCAG

```

FIG. 30

```

34501 CACACTGCTA CTTGGGTCAA TGACATTTTT ACTCCCTTGT TTTGTAGCTC
34551 AATTGGGTAT TACTGATGGG ATTTTGTAAT TATTAATATT TTCTTGTCTC
34601 CATTTTCTTC TCAAGTACTT TGTGCTTTT GAGTAAAATG CTTGCTAAGG
34651 GTATAGTTTT CACATAAAAG CTCAAATTTA GCATGGAAAT TAAGATATGC
34701 TCATACGTCT GCCATCCCTT ATCTGTAATT CTGAAATACC TAGAGTTCTG
34751 AATAACCTCA AATTCTTTTG TTACTTGTTT ATCAGCAAAA CCTGATTTGA
34801 ACTCAGTTTT TGGCAAAACT TGATCCAAGC TCTCTTAAGG CTCTTTTTAG
34851 TCTTTATTCA TTCCCTTTAG TGTGACTTCC CATTTTGCTA TAAAATTATG
34901 AGTGTGTTTG ATTACAAGGT GATGTCCCAG ACCCTACTGA GGGTGTTACA
34951 TAATATAAAC TGTATGTATG GCTGGGCGCG GTGGCTTATA CCTGTAATCC
35001 CAGCACTTTG GGAGGCCGAG GCGAGCGGAT AACCTTAGTT CAGGAGTTCA
35051 AGCCCAGCCT GGCCAACATG GTGAAACCCC GTCTCTACTA AGAATACAAA
35101 AATTAGCCAG GCATGATGGT GGGCGCCTGT AATCCCAGCT ACTCCTTAGG
35151 CTGAGGCAGG AGAATCACTT GAACCCAGGA GGTGGAGGTT GCAGTGAGCC
35201 AAGGTCATGC CACTGCACTC CAGCCTGGGC GACAAAGCAA GAATCTGTCT
35251 CAAAAAAAAA AAAAAAAAAA TGTGTGTACC ACTTTACCTT TCTAAAATCT
35301 GAAAAATTCT GAATCTGGAA ACCCATTTCTG CTTCAAGATA AATGGATCCT
35351 AGATTTATAT CGGTACCGTA CAGTCCTGAA ATTCCATCCT ATCTATTGGC
35401 CACTTTTACA TCAACAAACC TTTGAAGTTT GGGGAAACTT ACATATCACG
35451 CTCCCTTGGC AGTTGAACAT TATTTATTTA TTTTGAGATG GAGTCTCGCT
35501 TTGCCCAGGC TGGAGTGCAG TGGCGCGATC TTGGCTCACT GCAACCTCTG
35551 CCTCCCGGGT TCAAGCAATT CTCCTGCCTC AGCCTCCTGA GTAGCTGGGA
35601 TTATAGGCAT GCAACACCAT GCCCAGCTAA TTTTGTCTT TTAGTAGAG
35651 ACGGGGTTTC ACTATGTTAA CCAGGCTGTT CTCGAACTCC TGACCTTGTA
35701 ATCTTCCCTC CTCGGCCTCC CAAAGTGCTG GAATTACAGG CGTGAACCAC
35751 CACGCCTGGC CCTGAAGATA CATTTTAAAT CAATGAAAAA AACAACAGGA
35801 TTCTACCTCC TATGGTATAT CCCTCCTGGC TGTCTCTTCT CTCCAGTCTT
35851 GCCTCTGCTG TGTGGGTTTC AGGCATCCAT CTTCTCTACT CTGAATTACT
35901 GTGATAACCT CTGAAGTATT TTCCCTGCCA TCTGTCTGGC CCTTCTCCCA
35951 GGTCTTCCAC ATACTGCAGC CAAGTCAGCC CGCTGTTGAA ACCCTTCAAG
36001 ACTCCCTGCT GTCCTCTGGA TGAAGTCCAG ACTCTTCCAC GTGACTTACC
36051 AGGCCTTTCT TGCAC TTGTC CCCAGCCACT TACTGTTTCT CTCTTTCTAC
36101 CTTAACATCC TGAAC TTCTT TTGGTTCTTT GACCTTGCCCT CTGACCTTTT
36151 TCCATGCTGT TCACTCTTTC CCTGTTTACC TTGCTAACTC CTCTTTCTCT
36201 TTCTGGGTTG GATCAGATT CACTTCTTCC AGAAGCCCTT CCTAGACCTT
36251 ATACTTCTGG AATGGCGCCT TTTGACTGTA CGCTCATTGC ACCCTGTACT
36301 TCTCCTTTAT GAGTGGGTGC TGGTCTGTCC CACTAGGCTA CTTCATCCAT
36351 AAAGGGAGAG TAGAGCTTTA CCAAGTCAAT GCTTAAGCAA TATTTATTGG
36401 ATGAATGTGT GATTAATTTT ATAGAAATTT GATGTGCATT CAAATTTACT
36451 TATTGTATTA CGGAAC TTGTC ATTATATTCT CAGTGGAGTT ATTTTCTTTC
36501 ACGTGTGTAA TTCAAGATAG ACTCAGTGAG ATTTTCAAAA TTTGGAATGC
36551 AGTGCAAGGA AATTGAACTT GAGTTCTTTT GCATTTTGAT GGTTAAAAAT
36601 TTCCCATTTG TGGTGACATA CCACAATAAG CCAGTGAATG TGGCTTATTG
36651 TTTTCTGGTC TATAGAAAAT TGTCGCAAAC TCTGTCATAA TGTCTGGTTC
36701 TATATAACAA AGCTAGTCCT GTATTCTGCA TGTGGCTGAT GGAAACAGTG
36751 CTCTGTTGAT CTGGTTCATG AAGAAATCTG TTCAATTCTG CATAACAGAT

```

FIG. 3P


```

36801 GCCTTCATCA GTGTCCTTCC ATGAAGGAGC TGATCTTCAC AAAGAACACA
36851 TAGTTTTGCA TCCCACCACT TGCAGTATTT TTTTTTTTTT TTTTTTTTTT
36901 TTGAGATGCA GTCTCGCTCT GTCACCCTGG CTGGAGTGCA GTGGCATGAT
36951 CTCAGCTCAG TGCAACCTCT ACCTCCTGGG TTCAATTGAT TCTCCTGCCT
37001 CAGCCTCCTG AGTAGCTGGG ATTACAGGCG CACACCACCA TGCCTGGCTA
37051 ATTTTTGTTG TTTTAGTAGA GACGGAGTTT CACCATATTG GTCAGGCTGG
37101 TCTCAAACCTC TTGACCTCAT GATCTGCCTG CCTTGGCGTC CCAAAGTGTT
37151 GGGATTACAG GCGTGAGTCA CTGTGCCCTG CCAGTATTGT TTTGTCTAAA
37201 TTATTTGTGC TGATGTTTTT CCTACTGTGG TTTTCTTCAG ATTACCCTTG
37251 CTCTGAGCCT GCAATTGACT CATGAACCTC TTTTCCATGT TCTAACCTTA
37301 CAATGACTTC CTTGTGTTCA CTCCAAATGT TTTTCCCTGG TTGCATGTAG
37351 AGATGTATTA GCTAAGGTAC ATGCTTAGCT GCTGTATCAA AGAGACCCTA
37401 ATGTACAACC CAGGCTGGTA GAGCAGCTCT GCTGTATGTG TTAATTCAGG
37451 GACCCAGGTT CCTTCCATGT TGTGACTCCC CCCTTCCTTA GGATGTTGTC
37501 TTCTTTTACA TGGCTGAAGT TGGGCCATTT CATGTCTCTG TTCCAGCTGC
37551 CTGGTAGGAA AAAAGAACAG AAATTCAGAG TAAGCAAATT CTTTTTCTAT
37601 AGATGGATGC GGAAGTTGGA CACATCATTT CCTCTCACAT TTTCTCGGCC
37651 AGAACGTAGT CATGTGACTG CACGTCTAGC TGCTAAGGAG ACTGGGAATT
37701 TACTGTCGGC TGTGTGGCCT CTGTCAAGCT AAAATTCTTA TTACTGTGGA
37751 ATAAGGGAAG GATGGATTTG GGGGCACAAAT TAATAGTCTG TCACAGAGGC
37801 TAAAACAGCT GCTTTTGGCT GGGCACGGTG GCTCACACAT GTAATTTGAG
37851 CACTTTGGGA GGCCGAGGCA AGTGGATCAC TTGAGATCAG GAATTTGAGA
37901 CCAGCCTGGC CAACATGGTG AAACCCTGTC TCTCCTAAAA ATATAGAAAT
37951 TAGCCGGGCA TGGTGGCGGG TACCTGTAAT CCGAGCTACT CCAGAGGTTG
38001 AGGCAGGAGA ATTGCTTGAA CCTGGAAGGC AGAGGTTGCA GTGAGCCAAG
38051 ATGGTGCCAC TGCACCTCAG CCTGGGCGAC AGAGCAAGAC TCCATCTCAA
38101 AAAAAAAAAA AAAGGTTAAA TAAACAGCTG CTTTGTAGG TGATACAAGG
38151 TACAGCTAAG CTTTGAAGCC AGGCCTGTAG TTTCACCTTC CATATTCTTA
38201 CTCAAGGCAT TATACTTCTG GATCTGAAAC CACTGGATCT GATGCCCTGC
38251 TTGGGATGAG TTCTTTATAT TATCTTGCTT TCAACCCACA CCTGTGTAAT
38301 TTTATGGGCA GCGTTTGT TTCTATATAGG AACAATTTGA AAGTGGGCTG
38351 TTTCTAGGCT TTCATGAATA GCAGGCTATG CTGTCAATTGG GAATCTGGAG
38401 GGAGTTAATG AACACAACCTT CATTTGTTTAC TTTAGTGAAA TGTGGCAGCT
38451 TATGATAGTT TTGACAGTGA GACATGTGCT GTTTTGATCT CTCAGCTAAG
38501 ATTATCTGAT TTTTCAGGCA TGTCTCAAAA CTCACCAGGC CTGCTCACAT
38551 GCTGCTGCTT CTGAAGCCAG GGTTCGGAAA CCAGCTGCCC ATCAGAATGA
38601 GGCTGTGACT TAGAATATTG GTTCTTGTTT TATTACCATT CCTTGTTTGG
38651 TCTCTCCAGA GTCACCTGGC TTTTCCGCTT CAATTTTCTT ATCGGTGAAA
38701 TGAGATATTA ATTCCTCTTA TTGACTTCAA TTCAATTGCT GAGTGTATTG
38751 TTGCCTTTGG GAGGTTCTTT GAGTTTTCTG TGCCTTTGAA ATAGTTGTTT
38801 TTTTTTATTC TGGTGTTTTG AGGCATGTTT CAAGTGAGTG CATTTACACT
38851 TCTACCATTT TAGGAGCCAC AATTCAGTTA TGTTGTCCCA GCTTGCTTGG
38901 CCCCATCCCC AGAGTTTCTG ATTCAGTAGG TCTGGGGTGG GGCCCAATAA
38951 TTTGCATTTT TTCTTCTTTT TTCGAGACAG AGTCTGACTG TGTATCCAA
39001 GCTGGAGTGC AGTGGCACGA TCGTAGCTCA TTGTAGCCTC AAATCCTGG
39051 GCTCAAGCCG TCCTCCCACC TCACCCTCCT GAGTAGCTGG GACTATAGGC

```

FIG. 3Q

39101 ATATACTACC ATGCCCTGCC ACCTTTTTTAA TTTTTTGTAA GGATGGGGGT
39151 CTCACTGTGT TGCTCAGGCT GGTCTTGAAT TCCTGGGCTG AAGTGATCCT
39201 CCTGCTTCAG CCTCCCCAAA TGCCGGCATT CCTGGCATGA GCCACTGCAC
39251 TTGGCCAAGA CTTTGCATTT CTAAGTAGTT TCCAGGTAAT GCTGCTGCTG
39301 GTGTAGGGAC CTCATTTTGA GAACCATTTG TCTATAGCTG TAGCTATAGT
39351 TAGTTTCTGG TTATAGCTTC TTCTTTTGT CCCTTCAGTA ATAGTGTACA
39401 CATCCGAAAT CCCTGTCCTT GCTCTTTCAG GCCCAGGCAT GGTATCTGGT
39451 CCTCTTCTGT TGCTAGCCCT GGGGTGCTTC ATCATCCCAA GTTTATTTTT
39501 CTTCTCCTAA CCTGAACCTT TGTAAATAGC CCCTTCCCTA ATGAACGTCC
39551 TCAATTCCCT GTTTTGCGTG TCCTGTCTGT TTCTTGGCAA GACTCTGGAT
39601 GATTCAGTAC TCAATGAGGA TTTTTCGCAT AGATGGATGA AACAGGCTGG
39651 GTTTCATGTT TTCTAAGATA AAGGTGCTTC TCTCTTTTTC TCTTGGTCAC
39701 TTTGACCAAG AAGAAAATAA CAGAGTTTTT ATTCTCAAGA AGAATAATAT
39751 CGGGGCCACT CTGCTCAGAG GCCACTCTGC TTTGAGGACC CCTTCTCTCC
39801 TCCCTCATGC CAAAGATCAG GAACATTGGG CAGAGCGGAT AACGATGCCG
39851 CCAGCGTCAT TACATTTTCA CGGCACTTTC AGTTGTGCTG AGCGTGCAAA
39901 CATTTCAAGG AGACATTTCT AAGAGGTGGC TAGCACAGCA TGCTCTAAT
39951 GCCCTATGTG AATTGGAATA GAGTACTAAA GAACTGTTCA ATATTCACCC
40001 CATCCCCGCA TATGCAAGCA TGCACGTGGG TTCATTGTAT ATGTGTGTGT
40051 GCACGTGTGC ACAGACACAT TTGTCCTTCG TTTCAAATGC AACACAATGG
40101 ATGGAAATTG CCTTCCTGGT ACTGGGGTAT GGATGCAAAC ACCAACAGAG
40151 AAGCAGCCGC TACTTCCAAA CTGAACACAT GTGAGATTTG CCCTTTAATT
40201 AGCATCTGCA GCTGCTGCCA TCAGAAGGGT CTGTCTCTGT TGGCCTGAAA
40251 GTCTTTGCTT TAAAAGAGCA AGTCCATTAT AGCTCCAAGC CAGGCTCGTC
40301 TGTGAGCTGC TGTGCTTTCT CTGCCATCAG CGGGGTGCGC ACATTGTTTT
40351 GGGCTGTTTC ACTCTAGGAC TCTTTCCTCC TCCTGTGCCC CCAGCCTTTG
40401 ATTACCATGC CTTGGTGATC CTCATTTGGG TGACCTGCAG CTGCTCATTG
40451 TGTGTGCAGG AGACATCTCC AGTCCTTGTA AGGAGGGAAG ATCACTGGCT
40501 TCAGTGCTGA TGGACTGGTT ATTTTCCAGC CCTTTGTCGT CAGTGATCTT
40551 GTCTTGATAT GCAGAAAGGC TCCAGGTAGT CACTGAAAAA AATATAAGCA
40601 GCAGAGGTGA TGGCTATATG AAAGTCACGT TTCATCAAGG GCATTGCTGC
40651 TATGGAAACT TTCAATTCAC TTGGAGTAGG GAGCCATATT GGTTCACAG
40701 CCTCCTCAGC AGTGGGTCCC AACACAGTGC TGGGCTAGCT GCCTCTGAAT
40751 CACCGCAGTA GCTCCTTTTA CTATAGATTC CTGGGTCCCA CCCATGGAAT
40801 GTGATCCATG AAGTCTGGGG TTATTCCTTG GAATCCTTTA AGCTCCCTAA
40851 GTGGTTGGGA TGGGAAAGAG ATATGCTTTA TGTTACTATA CTTCTTCTTA
40901 TTATTATTTT AAAATTCTTG CCGGGCGCAG TGGCTCACAC CTGTAATCCC
40951 AGCACATTGG GAGACCGAGG CGGGTGGATC ACTTGAGGTC AGGAGTTCGA
41001 GACTGGCCTG GCCAACATGA TGAAATCCCG TCTCTACTAA AAATACAAAA
41051 ATTAGCTGGG CATGGTGGCG CATGCCTGTA GTCCCAGCCA CTCCGGAGGC
41101 TGAGGCAGGA GAATCGCTTG AACC CGGGAG GCAGAGGTTG CAGTGAGCCG
41151 AGATCGTGGC ACTGCACTCC AGCCTGGGTA ACAGAGTGAG ACTTCATCTC
41201 AAAAAAACC CAAAAAACA AAATCTTTT TCATTATACC GGAACGTCAG
41251 CTTTATGGAG TCGGGGATTT TTTCTGTTTT ATCACTGCT GTTCCCTAA
41301 CATCTAGAAT AGTGGCTGGC ACGATAGGCA CTCAAGTATT GATTTAGATG
41351 AGTCTATTTT ATTTCTTTT AAATTTTAA TTTTATTAG AGGTGGGGTC

FIG. 3R

```

41401 TGGCTTTGTT GCCCAAGCTG GTCTCAAAAC TCCTGGCCTC AAGCGATTGT
41451 ACTGCCTCAG CCTCCCAAAG GGCTAGGATA GGCATGAGCC AACATGCCTG
41501 GCTTGTCTTA TTTTAAACAA GCACTTCTGG TGATTCTGAT GGACAATCAG
41551 GCTTGGGAAG TTCTAACCTA GAGGACCTAC AGTTGTCTTG GGGTAGAAGC
41601 CAAGGCTATC CTGGTTTTTA GAATCAGTGC CTTACTGGGC ATCTCTGAAG
41651 AGTAAAAGTC AGGGACAGAG TTACATTTTT GGACAAAACC AGATGCTGTG
41701 AATGGACTCT TGGTCACAAC CTGGGTGGCG ACTTGGTCCT TAACTTCTTC
41751 ATCATTTTCT GCTGACCCTG TTCTTTGGTT CACAGCAAGT CACCTGATAA
41801 GAAGACTCAA AGACTGCTAG TTTGTTACTT TAGATGATGC TTTTGAACCC
41851 TCTTGGTACC ATTTTAACAA TCCAAACGTA TTTTATGAAA GCACTCAAGT
41901 CCTGGGTCTT TATTGTATCT TTAAGCTCTA ACAGCATGAT GATTGAATAA
41951 GCTGTGGTTG GCCACACACA AGCCATCTTC CCCATGGCCT CCATTCATAC
42001 TAGAATGAGC AGCTATACCC CAGTAGTATA GTTTTGGGAT ATGGGTAAACA
42051 TCTTGGGATA GCCACATTTA CTTAGTAAAT GTCTGGCTTA CATTCTCCTA
42101 ATGGTGCAC TTTGGAATTT TTGGTGTGGT AACCTGGAAT AGTGTGGTG
42151 GGTCAAGTTT GATTAGCATC TTTGATAAGG ACCCGGTCTA TTTAGAGGTT
42201 TGTCAATTGAG TGTGTCTGTT TTGGCCTCAT GTTGTGAAGC ATGCTGTGTA
42251 GCAGCTGTTG TAATTTTGTG TGCTTGTTTT CTCAATCAAC CCTGGTTTTG
42301 AAGAAATGGG AAGTTGTTCC ACTCTTAGAC TGATCTGACT TGGGAGGGGA
42351 TTTTCAGTTC AGGAAGTTGG ATCTTCTGAA TGGAAGCAAA GAATACATGT
42401 CTTTTTGCCA CTTTACAAGC TGGCTCTTGT TTTCTGAACT ATTTTACTGG
42451 TCATTGCAAA TAGAATGTCA GGAGTAGCTG CCAAATACTA AGTTGTGTTT
42501 AGTTTGTGAG TTCTTAAGAG TTGCCGGTGG CTGCTCTGCT ATGCGTATGA
42551 CTTTCTCAGC CTTAAACTTA CAAGCCATAC TGTTTTTTTC ACATCTTTAA
42601 TACAGCCATA GGAAATTTAT AACTGTGGCG TGTCGTCATA AATATGCATT
42651 GTTCTTATTT TAAGACATTT CAGTACTAAA AGTATAAGTA CTTCTGTTAT
42701 TATCTGTGAA TTTCTTTCCT TCTTCTTTTT TTGGATATTT AAGACCTTTT
42751 CGATGTCAAT ATATATTTAA AACAGACATA TAAATTAGCA TTCACCCACA
42801 TACCCAGGGC CTATGGAGAA CCAGGTTGGG ATGAGTGGGT GAGCTACAGG
42851 CAGCCAGGTG GCTCCTGTGG GCTCCTCGAG GACTGGGGTG AGTAACTAAT
42901 GTCTGCTAGG AACTTGGGGG AAAGAAGGTG TGTATGTTAG GTGCTGCCCC
42951 CTTCTAAGTG TTCCTCTTGT TCATAATTTT TTTTTTTTTT TTTTTTTTTA
43001 GATGGAGTCT CGCTCTGTTG CCAGGCTGGA GTGCAGTGGT GTGATCTCAG
43051 CTCAGTCAA CCTCTGCCTC CCGGGTTCAA GTGATTCTCC TGCCTCAGCC
43101 TCCCAGTAG CTGGGACTAC AGGCATGCAC CACCATGCCC AGCTAATTTT
43151 TGTATTTTTA GTAAAGACGG GGTTTCACCA TGTTGGCCAG GGTGGTCTCG
43201 ATCTCTTGAC CTTGTGATCC GCCTGCCTCG GCCTCCCAA GTGCTGGGAT
43251 TACAGGTGTG AGCCACTGTG CCCAGCCCAT AAATCAAAAT TTTTTCAGCA
43301 ATTGTTATAC AAGTGAACCC TTACTCTTCA AATGCAATTG TCCAGTGTCT
43351 GGCTTAATGT CTGCTGTTGT CAGAAACCAT GTGAATGGAG TAGATTCCCA
43401 GGTTATAAGG AGCCCCCAGG GAGGATGCGC GAGTCACTGG CTTCTCCAGG
43451 GGTCTCTGGT TTGGGGTTGC CTTGGTGCTG GGCACACTTC CTGGAGATTT
43501 TACTGGACCA GCCTGAGGCC TTTGGGGCTC TGTGCAGATG CTCTACTTCT
43551 GACTTGTCTA GAGCTTTCTT CTAATTCTGG ACTAAAAGCA AGCAGGAGTT
43601 TGGAGGATGA TGGTGAGAAT TCACATCCCC GAGTTGGCTT TTGGAATGCA
43651 GTAGTTTGTG AGATTTAGTG TTTTTTTTAA GAAGTATATT CAGATCTTGC

```

FIG. 3S

```

43701 CTTTTTCCCA GAAAGCATAT GAGACAACCTT CCAAGACATT TATAGCATGG
43751 CTAATAAAAT GGGAAATCAG GGCGAAGGAC AGGAGAACTC AATAAGGGTT
43801 AACATGGCTA CAGCGATTGT CTAAATGGGT TCTTTTTGCT GGCCAGAGCA
43851 GAAAGGATCA TGCAGTAAAG TGGGGGGGAA GAAAGGGAAT TGAATGGTAG
43901 GTGAAGACTT CATGTTGGTG CCAGGCACTG TGCCAGGCCC TCCTAGGACC
43951 TTGTCTTACT CAATCCTCAC ACAGTGCTGC AAGAGGATTA GTCTTATCCC
44001 TGTTTTAGAG AGGATGAAAC TGAAAGGCAG CGAGGTGAAG TCACCAGCAG
44051 GAGGCTGAAG CCGCCCAGGC TAACTGGCCT TATAGCTACC TAGGGACTCA
44101 GGAATATCAC ACCTGTTTAT CATCAAAAGG AGAAAGGATT TCAGTTCCTT
44151 GGGGTAGAAG AGTTTCTTTT TGCTAATCAA ACATTTTACT TGAGGCTTCA
44201 TATTCTTCTT CAAGATTTTT TTCCTGTGTA TGTACCAACA CATGTAATAA
44251 TTCCTTGTTT ATTTCAAAAA AGGGGTGTGA CTTTATTCTT TACAAGATTT
44301 CACTTTATAT TGTCAATGGAC AATTTTCCAT GGCAGTATGA ATAAATGGAA
44351 TCTGTTTGTT TTTAATATCT TTGTCTTATC CCATTGTTTA CATATGTCAT
44401 ATTTTAGCCA GTCTCTAACT GATGGATAGC TGAATGATTT CCATGTTTTT
44451 TTCCCCTGTT ACAAACAATA CTGCAAGGAA TCTATTTATC TTTCTATTTA
44501 TCTGCAAACT ATTGTAAGTA CCTGTAAATT GTTAGAAGTG GAATTACTAG
44551 GTCAAAGGGG ATATTTTCAC ATTTAAATTT TGAATAGAGG CTGTCAGTTG
44601 CCTTCCACAC TGAATAATAA AGGAAAAGAT TGTATCACAT TTATTGCAAG
44651 CCTTCTGTAT TCTGCTGGGT GCTGAGGGGA ATACAGAAAG GATATAAGAG
44701 TGGTTGCCCT CTAGGAATAT CCGTCTACAC TGTACCTAAT CCTAGGGAAT
44751 GTCTGGGGTG TCAACTTGTG GGTGGGAAAG TGGGTGGATT TAATTCAACT
44801 GTTCAAGCTT GCCTTGCAAA CACTGTGCAT GGTGTCTGGG ACTAGTCTTT
44851 CATTATATTG ATTCCCCTGG GTAACAGATG TAATTTCTT AGGGCAGGGA
44901 CTTATCCTTA CATGACTTAC AGCGTGCCTT ACACATCTTC TTTGCTTTGT
44951 GGAGACCTTG TTATTATAAC ACGTCAGGTG ATATTCGAGG ATCTAATTGA
45001 GGCATTCCCT ATTTTTGGGT GTGTGAAGAA TTAATAACTT TGGCATTCTA
45051 TACAGGTCAT GGAATATCAG CCTGGAGGGG ACTTGCTGTC ACTTTTGAAT
45101 AGATATGAGG ACCAGTTAGA TGAAAACCTG ATACAGTTTT ACCTAGCTGA
45151 GCTGATTTTG GCTGTTTACA GCGTTCATCT GATGGGATAC GTGCATCGGT
45201 AAGTGAGACT CTGGTAGCAT TTTTATGCTG AGGATTTTCC TGTGTCGCAT
45251 AAGAGTTCCT GCATGGAAAT GAGTGGATGA GTGATTTCAA GATCAAGATA
45301 ACGCCCCATC CAGTTTTTAG CCAGTCTACC AATAACTGGC TGAAAGCAAA
45351 CTTTCCAAGA TGGAGGACAT TTCAGCTTGC TTATCCAGCA GTGCAATAGA
45401 TCTAGAATTG TAATGTGCTC AAGTTTGCTA GTAATATCTA TTAATGTAGC
45451 TAAATAAGAC TGGGAACTCT TGCATGGGTT CTTTGGGTTA TATGATAGAA
45501 GAACTGAATT TGGTTTGCAG AAGGAAATGT CATACCACAT AGTAGTGTA
45551 GACCATGGAG CTGTACTTCT CTAACCTCTGC CCGTTAGAAAT TTACAATTTT
45601 TTTTTTTTTT TTTTTTTTTG AGACAGAGTC TAGCTCTGTT GCCAGGCTGG
45651 AGTGCAGTGG TACCATATTG GCTCATGGCA ACCTCCGCCT CCTGGGTTCA
45701 GGTGATTCTC CTGCCTCAGC CTCCCAAGTA GCTGGAATTA CAGGCACGCA
45751 CCGCCATGCC CAGCTAATTT TTGTATTTTT AGTAGAGATG GGGTTTCACC
45801 AGGTTGGCCA GGATGGTCTT GATCTCCTGA CCTCATGATC CACCCACCTT
45851 GTCCTCCCAA ATTGCTGGGA TTACAGGCAT GAGCCACCAT GCCTGGCCTA
45901 CAAAATCCTC AGTTGGTAAG TGGTCTTCA TGTCTTCATT CATCTGATGT
45951 TTTGTGTACA TCTGAGAATG TTGTGGGAAT ACAATGATTG TTAGTCCAGG

```

FIG. 3T

```

46001 AATCACAAAA TTTGAGATAG AGTCTCAGCT TTTCCATTGC CTAGCTACAT
46051 GACCTTGGGA AAATTTTCATA GTCCTTTTGG GCCTTAGTTT TCCTCATGTG
46101 AAATGTGTGT CTCTAGGAGA AATAATCCAT TGAATAATAT GTGTTTCATT
46151 TCTCTTCCTT TTCTTTCTCT CCTATCCTTC CTGCTCCCT CTCGCCCTTT
46201 TTCTCTTTCC CCCTCTCTCC CTCTCTCTCT CCTTCCTTCC TTCCTTTCGG
46251 TTAAATTCAT TTTGCAAAAT GTATGCTAAT AATTTATATC CACCAATAGA
46301 GGAGGTCTAT ATAACAGAAT ACATAAACAA AGATTTTTTGG CTCAATTGAG
46351 ATTCTAGGTT AGCACTTGCT TGCTGATTGG GATGGAGGAG GCAATTCATG
46401 GTCCTGATTT TCTTACAGAG ACATCAAGCC TGAGAACATT CTCGTTGACC
46451 GCACAGGACA CATCAAGCTG GTGGATTTTG GATCTGCCGC GAAAATGAAT
46501 TCAAACAAGA TGGTAAAAAA TGGAATAAGA TAGCTTAATA GAGTTTATAC
46551 TAAAAAGTAT TCTTGGTCCT CCTAAGTTTG GGAAGTGTTG GGATAAAATG
46601 GTGAACAATG TTTTGGAGCC TTTGGCAGTG TATGGGGGTG GGGACAGGGA
46651 CACAGAACCA TTTCCCAGAC CGTGGCACCT TTTTATTTAT AGTGCCTGTT
46701 AATACCCTCC AAGACATTTT TAGGAGCATT GTTATAGTTT GGTAGAAAT
46751 AAAGGAAAT GCTTATTTTG TTTCTCTCTT CATTTTCCTT GCCTGTTATA
46801 GACTGTCTTT TGTATATTA TCTTTTTTAC TTTAAATAT TTTGATGAAA
46851 TGGAACTCC TGCATGTCAA ATCCTCTATT TCCTATGCAG CAAAATTGAA
46901 ATTAATCACT GGAGCATTTG AACCAAATAT CCTTAAGTGT TAAGAACCAA
46951 GTGCTCAAAA TATCATTTTT AAGTCTTGA TCTTTGGTAG AAATTAACT
47001 GTATTCACA TGCTAAGTAG GACGGCAGGA GGGTAGCTAC TGAGATCAAG
47051 AGTGAGACTA CTTTAGGAAA AAGATGACAA AGTAAAAAAA GATTAGAGTT
47101 TAAAAATCTT CTAATAAAGT TGGTATGTAC TAAAATATGA ATTTGGAAGT
47151 CAACTCCGCA AAAAAGGATA GGTCTAAGAG AAAATCGACT TAGGTTTTAA
47201 GACTGATTTT ACAACTGAGC CATTTGGTGA CCTAGACAAA TCCTTGGGAA
47251 CTTGATCTTT TATACTTTCT CTAGAAAAAA CTGATGCTAG TGAAAATGCA
47301 TAATTTAAGA GGTTAGAGAA GCTGCTCTTC AAAATGCCCC CCAAGTCTGA
47351 GAGTTAAATC CTTTACATAA AGGACAATAT GTAAAATTTT CTTTTTCTTT
47401 TTTCTTTTTT TTTGAGACGG AGTCTCGCTC TGTCCTCCAG GCTGGAGTGC
47451 AGTGGCGCGA TCTCGGCTCA CTGCAAGCTC CGCCCCCTG GTTTCACGCC
47501 ATTCTCCTGC CTCAGCCTCC CGAGTAGCTG GGACTGCAAG CGCCCGCCAC
47551 CATGCCCAGC TAATTTTTTG TATTTTTAGT AGAGACGGGG TTTCACCGTG
47601 TTAGCCAGGA TGGTCTCGAT CTCCTGACCT CGTGATCCAC TCGCTTCGGC
47651 CTCCCAAAGT GCTGGGATTA CAGGCATAAG CCACTGCGCC CGGCTCTTTT
47701 TTTCTTAAA CTGCTTCCAG AAAAGTGGAT ATTATTAGGT TGATGTTAAG
47751 AAAAGGCTTG GAGTTGCATT AACTTTTTGC TTTCTAGCAT CTGGCCTGTC
47801 TGTTCTGCAG ACCTGAGACC TACTTGAGAT AATTTTCTTG GTGTTCAAGC
47851 CCTTGAAAA ATAAGTTCCC TATGTTGTCC AGTGTCAAAG TTTCTCAACC
47901 TCAGCACTAT TCTTTTTTTC AGGTTATTTT CTTGTAATCT GTTCACTTGA
47951 TCATTACATT AAGAATTAGA TTATATTGCT ATAACATAA AGCATTTTAT
48001 GTTTTAAAA TTATGTACAA TTTAGAAACA GGCATGAAAA CTTAGGTATT
48051 AAATTTAGTG GAATAAAGCA CAGAAAAAAA GTTAAATAA TGCAGTTTAA
48101 TCACTTAGGA TTAAACATTT ATATGGGCCG GGTGTAGTGC CTCACACCTG
48151 TAATCCCAGC ACGTTTGGAG GTCGAGGCGG GAGGATTGCT GGAGTTTGAG
48201 ACCAGCCTGG GCAACAAAT GAGACCTAGT CTCTACAAAA AATCAAAAAA
48251 TTAGCCAGAC ATGGTAGTAC ATGCTTGTAG CTCCAGCCAC ATGGGAGGCC

```

FIG. 3U

48301 AAGACAGTAG GATCGCTGGA GCGAAGGAGG TTGAGGCTGC AATGACCGTG
48351 TTTGCACCAT TGCATTCCAG CCTGGGCGAC AGAACAAGAC CCTGTCTTAA
48401 AACAAATTTA TATGCTGCAT TCGTGAAATT AAAAAAAAAAT CATGGATTTA
48451 GAAATAAATT GAAGCAAGGT ACATTGACAG TGTAACCTCA GCACTACTGA
48501 CATTTTGATC TGAATAATTC TTTGTTGTGG GGGATGCGCT GTATAAGATG
48551 TTTAGCTGCA TCCCTGACTC CTACCTCCTA GATGCCATTA GCACCCTCCC
48601 CTCCAGATGT GATAACCAAA AATGTCTCTA GACATTGCCA GATGTGCCTG
48651 GGGTAGGAGG GTTGGGGGAA GTGGGGTTTG AGAACCCTTA GTTGATCATG
48701 CCTGCAGTAG GTTGAGAAGC ATCAGAAAGC TAATTAATTA GACAGGAATA
48751 TGTGTTTGCA GTA (SEQ ID NO:3)

FEATURES:

Genewise results:

Start: 3121
Exon: 3121-3216
Exon: 10089-10230
Exon: 21592-21767
Exon: 29016-29117
Exon: 45056-45198
Exon: 46419-46530
Stop: 46531

Sim4 results:

Exon: 2001-2040, (Transcript Position: 1-40)
Exon: 3108-3216, (Transcript Position: 41-149)
Exon: 10089-10230, (Transcript Position: 150-291)
Exon: 21592-21767, (Transcript Position: 292-467)
Exon: 29016-29117, (Transcript Position: 468-569)
Exon: 45056-45198, (Transcript Position: 570-712)
Exon: 46419-46764, (Transcript Position: 713-1058)

CHROMOSOME MAP POSITION:

Chromosome 12

ALLELIC VARIANTS (SNPs):

DNA			
Position	Major	Minor	Domain
722	A	G	Intron
11380	T	- A	Intron
14282	G	C	Intron
30482	A	G	Intron
30903	G	A	Intron
31969	T	C	Intron
33307	G	- T	Intron
38763	A	G	Intron

FIG. 3V

38854	A	G	Intron
46559	G	A	Exon, 3' UTR
47193	G	A	Intron
48129	T	C	Intron
48676	G	-	Intron

Context:

DNA

Position

722	TCTGGCTGTTGCTTTCTTTATGGTTTTGTCATTACTTTAAACAATGACAAAACTGCAAT GATTTGCATCAACCTAATACATCCCTCCTTAAACAATGTTGCTTTGTTTTGTCCTGTTTT GGAACCTATAAGAATGGAATCATAATGGAATCATATGTTATTTTTCTTGCTTCCTTCATTA GGCCTTGTTTTGAGACTCATTATGTCATTGTGGTTAGTTGCAGTTTATTCTTTTTTCATTG CTTGTGAAAACACTGCAATATACAATTTTGTCTTTTCTACTGCTGATGGACATTTATATC [A, G] CTTCCAGTTTTTTTGCGAACACTATTTTGTATTCTTATACACATCTCTTGGTGTACATAAG TAGGAGTTTTCTCGCCGGCGTGGTGGCTCAGGGCCTGTAATCTCAGCACTTTGGGAGGCCG AGGTGGGCAGATCACTCGAGGTCAAGAGTTCAAGACCAGCCTGGCCAACACGGTGAAACC CCATCTCTACTAAAAATACAAACAATTGGGCATGGTGGCATGCACCTGTAATCCCAGTTA CTTGGGAGGATGAGACAAGAGAATAGCTTGAACCTGGGAGGTGGAGGTTGCAGTGAGGCG
11380	TTGCCCTATTACCCAGGCTGGATTGCAGTGGTATATCATGGCTCACTGCAGTTTCAACCT TCTAGGCTCAAGCAATCCTTCCACCCAGTGGCTGGGACTACAGGCTCACACTACCACGC CCAGCTAATTTTTGCTTTTTTCTCTGTAGAGATAGGGTCTTACTATGTTACCCAGGCTGG TTTCAAACCTCCAGGCTTGAAGCAGTCTTCTGCCTCAGCCTCCCAAAGCTTTGGGATTAC AGGTGTGAGCCACCATGCCTGGCCCCATAAAATATAATTTTTGAATTCTTTTTTGTTTTT [T, -, A] ATGGAGGAAGGGGCTGAGGAAGGCAAAAGTACCTAGGGCCTATGAAGTCATATATTGGCC TTGCCCTTCAACCTGTTTCTGACTTTGCTTGACTTCCATGTGATGAGGCAGTTGGCTGTTA GTGTCCCAGTTTCATACTCTTACATTAGTGTTTTTCAACCAGTGGGTGATTTGACGTTTT CGGTTGTGAGAGCTAGTTGGGGGTGGTGGTGTGTGAGTTTGGGGGAAGGGTCTACTGT CAGTTAATGGGTGAGGCCAGAGATGCCACCAACACCTTACAGTGCACAAAGCAGCCCCC
14282	CTTACTGAGCCATCTGCAGGCACCTTCATTAGTCTTGAGACTGTCCTCTGGTTACTTAAC AGCAGTGAATTATCTAGAATCATTTAGTGATCAGAAGACTTGGTTTAGTGGAATGTAGAT TTTTTCTAATAGACCCCTCTTCCAGGGAAATGTTTCATATTTTTGAAGAGGTTTCCTGG GGAGTGTTTAAGAGGCCATGATTGAAAATGGGTGATTACATTAGTGTGTTTTCTATTCCT CCCCTTTTTGAGTTTCTGTTTTGGAATGTAAGCTTTGTTTTTCTACGTGGAGAAGGGTCC [G, C] TCAGCTGCTTCTGCCAGGTTTTTTGAATCTTCTATAGGGATGGAGATTTTCTTTGGGG ACTGTTAGAGAAAATGGAATAGAGTGTAGCTCTGAAGGAGAAGGATGTCTCCAGCAGAAG TACCTCTAGCCTTGGGCCAAGGGAGGGAAGGGAAGGGAACGAGCATCTGGGAACCAGGGA AGGGATTTTTGTCTTTCTTAATTACTCTTACATCCCCAGTGCCCAAATAGTGTCTGGCA TATGTTAAGTCCTTAGTAAATACTTGTTGAATGAGTGTATGCTCAGTGAACAAAATAAAT
30482	AAGAAATGTTAAATGGTTACATAAGCTTTCCCTTTCTGACCCTTAAGTGTGCTCTGTAGG AGCATGGTGGGGGATGTTTCTTTTCTTTTCTTTTCTTTTGTGAGACCAGGTCTCACTTTGC CAGGAGGCTGGAGTTCAGTGGCATGAACATGGCTCACTGCAGCCTCGACTTCTGGGCT CCAGCAAACCTCCACCTCAGCCTCCCGGCATACACCACTGTGCCTGGCTAATTTTTGT

FIG. 3W

ATTTTGTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGGCTGGTTTTCGAAGTCCTGAGCTC
[A, G]
AGAGATCTTCCTGCCTTGGCCTTCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATGCC
CAGCTCCGGTGGGGGATATTTCTATATCCACATGTGTATAGTTTACTTTATAAAAAATGGT
ATGTTACTCTGTGCTTGGCTCTCCAGCTTGCTGTTGCCTTTTACCAGTGTATCCCAGACA
TCCTTTCTTCCTTGTCTAGTAACGCAGGTCTACTTTATTCTTTGAGCAGTGGCATAATTTT
CCCTGATGTGTATATATCATAAGTTAGAGAATGCTAAAATTCATTTTGGGGCCTTGTTTA

30903 ATGTTACTCTGTGCTTGGCTCTCCAGCTTGCTGTTGCCTTTTACCAGTGTATCCCAGACA
TCCTTTCTTCCTTGTCTAGTAACGCAGGTCTACTTTATTCTTTGAGCAGTGGCATAATTTT
CCCTGATGTGTATATATCATAAGTTAGAGAATGCTAAAATTCATTTTGGGGCCTTGTTTA
GGTTCTTGAGGGATTAAATTCCTAAATTTAACAAGTGTATCCTGGAAACAATTTTGTTC
CTGATTACAGCCCTTAAAAGAGGACTATCATGTTACCTTGAATGGAGATAAACAGGCTCAC
[G, A]
TAAGAGAAAAGGGTAAGAGGGATGAACTCCCACTTATCTTAACTTCTACTGGCCCCGTTT
TTGGGGAATTTGCTGCTTTTATTCCTGACCTAAAATAAATAAGTTTATGTGTCTTGGTTT
CATATTAGTTGAGAACCCAGTGCCCTGGAGAGAAGTTTCTCTGTCTCTGAGTGAGGACA
TTCACATATGAATCTATTGGCAGACTGGCTTTGACTGACCACACGTGCCTTCAGAACCAA
TGCCACAGCTCTTAGGTTTATGGCCTGAAACACCCCTTCTCTTACATATTGCCTTAGAAAC

31969 CACTGCGCCTTTTCTTTTCATTTTTTTCTGAGATGGAGTCTTCTCTGTACCAGGCTG
GAGTACAGTCATGCAATCTCAGCTCACTGCAACTTCCACCTCCTGGGTAAAGTGATTCT
CCTGTCTTAGCCTCCTGTGTAGCTGGGACTACAGGCGTGTGCCACCGTGCCAGCTAATT
TTTATATTTTGTAGTAGAGACGGGGTTTTGCCATGTGGGTAGGCTGGTCTTGAACCTCTG
ACCTCAGGTGATCCACCCGTCTTGGCCTCCCAAAGTGCTGGGGTTATAGGCGTGAGCCAC
[T, C]
GTGCCCAGCCTCAGGCTTCTTTATTAAGAAGAAGTTCGGGCCAGGTGTGGTGGCTTACAC
CTGTAATCCAGCAATTTGGGAGGCCGAGGTGGGAGATCAGGAGGTGAGGATCGAGA
CCATCCTGGCTAACATGGTGAAACCTCGTCTCTACTAAAAATATAAAAAATTAGGCAGGT
ATGGTGGCGGGTGCTGTAGTCCCAGCTACTCGGGAGGCTGAGGGAGGAGAACGGTGTGA
ACCTGGGAGGCGGAGCTTGCAGTGAGCCAGATTGTGCCAGTGCACTCCAGCCTGGGTGA

33307 TGAAGGAACTCAGTTGAGGATTTAGTGGTCCATGTGGCAGGCTACTGCTCAGTAGCTTCA
GGGAAACCACTGCTTGCCCTCCCCTGTGGCCAGTGAGGATGATCAGAGGAGTCCCAGCAGG
AATGCCCAAATGTAGTTTTCTTACATGTTGATGGGAGTGCATTGTTTCATGTCTAAACAG
TTCTCAAATCACATCTTCAGGAGGGTACTATCTGGGCACTTTGATAATTTCTCACTTTGA
TGTACCCGTTCTTATTACCATCACCTAGTTTTGTCTAGTAGAAATAACTTTCTTTTTTC
[G, -, T]
GTGAGATGGAGTCTTGCCG
TGTTGCCCAGGCTGTAGTGCAGTGGCGTGTCTCGGCTCACTGCAACCTCTGCCTCCCGG
GTTCTCCTGCCTCAGCCTCCCGAGTAGTTGGGATTACAGGCGTGTGACACCACGCCCGGC
TCATTTTTGTATTTTCAGTAGAGATGGGGTTTCACTACTTTGGCCAGGCTGGTCTTGAAC
TCCTGACCTTGTGATCCGCCACCTTGACCTCCCAAAGTGCTGGGATTGCAGGTGTGAGC

38763 GACAGTGAGACATGTGCTGTTTTGATCTCTCAGCTAAGATTATCTGATTTTTTTCAGGCATG
TCTCAAACTCACCAGGCCTGCTCACATGCTGCTGCTTCTGAAGCCAGGGTTTGGAAACC
AGCTGCCCATCAGAATGAGGCTGTGACTTAGAATATTGGTTCTTGTGTTTTATTACATTCC
TTGTTTGGTCTCTCCAGAGTCACTGGCCTTTTCCGCTTCAATTTCTTATCGGTGAAATG
AGATATTAAATCCTCTTATTGACTTCAATTCAATTGCTGAGTGTATTGTTGCCTTTGGGA
[A, G]

FIG. 3X

GTTCTTTGAGTTTTCTGTGCCTTTGAAATAGTTGTTTTTTTTTATTCTGGTGTTTTGAGG
CATGTTTCAAGTGAGTGCATTTACACTTCTACCATTTTAGGAGCCACAATTCAGTTATGT
TGTCCCAGCTTGCTTGGCCCCATCCCAGAGTTTCTGATTGAGTGGTCTGGGGTGGGGC
CCAATAATTTGCATTTCTTCTTCTTTTTTCGAGACAGAGTCTGACTGTGTATCCAAGCT
GGAGTGCAGTGGCACGATCGTAGCTCATTGTAGCCTCAAACCTCTGGGCTCAAGCCGTCC

38854 GCTGCTTCTGAAGCCAGGGTTTGGAAACCAGCTGCCCATCAGAATGAGGCTGTGACTTAG
AATATTGGTTCTTGTTTTTATTACCATTCCTTGTGGTCTCTCCAGAGTCACTGGCCTTT
TCCGCTTCAATTTTCTTATCGGTGAAATGAGATATTAATTCCTCTTATTGACTTCAATTC
AATTGCTGAGTGTATTGTTGCCCTTTGGGAGGTTCTTTGAGTTTTCTGTGCCTTTGAAATA
GTTGTTTTTTTTTATTCTGGTGTTTTGAGGCATGTTTCAAGTGAGTGCATTTACACTTCT
[A, G]
CCATTTTAGGAGCCACAATTCAGTTATGTTGTCCCAGCTTGCTTGGCCCCATCCCAGAG
TTTCTGATTGAGTGGTCTGGGGTGGGGCCCAATAATTTGCATTTCTTCTTTTTTTCG
AGACAGAGTCTGACTGTGTATCCAAGCTGGAGTGCAGTGGCACGATCGTAGCTCATTGT
AGCCTCAAACCTCTGGGCTCAAGCCGTCTCCACCTCACCTCCTGAGTAGCTGGGACT
ATAGGCATATACTACCATGCCCTGCCACCTTTTTAATTTTTTGTAAAGATGGGGTCTCA

46559 ATTTTGCAAAATGTATGCTAATAATTTATATCCACCAATAGAGGAGGTCTATATAACAGA
ATACATAAAACAAAGATTTTTGGCTCAATTGAGATTCTAGGTTAGCACTTGCTTGCTGATT
GGGATGGAGGAGGCAATTCATGGTCCGTGATTTTCTTACAGAGACATCAAGCCTGAGAACA
TTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTTGGATCTGCCGCGAAAATGA
ATTCAAACAAGATGGTAAAAAATGGAATAAGATAGCTTAATAGAGTTTATACTAAAAAGT
[G, A]
TTCTTGGTCCCTCCTAAGTTTGGGAAGTGTGGGATAAAATGGTGAACAATGTTTTGGAGC
CTTTGGCAGTGTATGGGGTGGGGACAGGGACACAGAACCATTTCCAGACCGTGGCACC
TTTTTATTTATAGTGCCTGTTAATACCTCCAAGACATTTTAGGAGCATTGTTATAGTT
TGGTTAGAAATAAAGGAAAATGCCTATTTTGTCTCTCTTCAATTTTCTTGCCTGTTAT
AGACTGTCTTTTGTATATTATCTTTTTTACTTTAAATATTTTGATGAAATGGAACCTC

47193 AAATTGAAATTAATCACTGGAGCATTTGAACCAAATATCCTTAAGTGTTAAGAACCAAGT
GCTCAAAATATCATTTTTAAGTCTTGGATCTTTGGTAGAAATTAAACTGTATCCACATG
CTAAGTAGGACGGCAGGAGGGTAGCTACTGAGATCAAGAGTGAGACTACTTTAGGAAAA
GATGACAAAGTAAAAAAGATTAGAGTTTAAAAATCTTCTAATAAAGTTGGTATGTACTA
AAATATGAATTTGGAAGTCAACTCCGCAAAAAGGATAGGTCTAAGAGAAAATCGACTTA
[G, A]
GTTTTAAGACTGATTTTACAACCTGAGCCATTTGGTGACCTAGACAAATCCTTGGGAACTT
GATCTTTTATACTTTCTCTAGAAAAACTGATGCTAGTGAAAATGCATAATTTAAGAGGT
TAGAGAAGCTGCTCTTCAAAATGCCCCCAAGTCTGAGAGTTAAATCCTTTACATAAAGG
ACAATATGTAAATTTTCTTTTCTTTTTCTTTTTTTTTGAGACGGAGTCTCGCTCTGT
CCCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCCCCCTGGGT

48129 ATAATTTTCTTGGTGTTTCAAGGCCCTTGGAAAAATAAGTTCCCTATGTTGTCCAGTGTCAA
AGTTTCTCAACCTCAGCACTATTCTTTTTTTTCAAGTTATTTTCTTGTAACTGTTCACCTT
GATCATTACATTAAGAATTAGATTATATTGCTATAACTACAAAGCATTTTATGTTTTAAA
AATTATGTACAATTTAGAAACAGGCATGAAAACCTAGGTATTAATTTAGTGAATAAAG
CACAGAAAAAAGTTAAAAATAATGCAGTTTTATCACTTAGGATTAAACATTTATATGGGC
[T, C]
GGGTGTAGTGCCTCACACCTGTAATCCCAGCACGTTTGGAGGTGAGGCGGGAGGATTGC
TGGAGTTTGTAGACCAGCCTGGGCAACAAAATGAGACCTAGTCTCTACAAAAAATCAAAA

FIG. 3Y

ATTAGCCAGACATGGTAGTACATGCTTGTAGCTCCAGCCACATGGGAGGCCAAGACAGTA
GGATCGCTGGAGCGAAGGAGGTTGAGGCTGCAATGACCGTGTTTGCACCATTGCATTCCA
GCCTGGGCGACAGAACAAAGACCCTGTCTTAAACAAATTTATATGCTGCATTTCGTGAAAT
48676 GCGACAGAACAAGACCCTGTCTTAAACAAATTTATATGCTGCATTTCGTGAAATTAAAAA
AAAATCATGGATTTAGAAATAAATTGAAGCAAGGTACATTGACAGTGTAACCTCAGCACT
ACTGACATTTTGATCTGAATAATTCTTTGTTGTGGGGGATGCGCTGTATAAGATGTTTAG
CTGCATCCCTGACTCCTACCTCCTAGATGCCATTAGCACCCCTCCCCTCCAGATGTGATAA
CCAAAAATGTCTCTAGACATTGCCAGATGTGCCTGGGGTAGGAGGGTTGGGGGAAGTGGG
[G, -]
TTTGAGAACCCTTAGTTGATCATGCCTGCAGTAGGTTGAGAAGCATCAGAAAGCTAATTA
ATTAGACAGGAATATGTGTTTGCAGTA

FIG. 3Z